The Economic Effects of Counterfeiting and Piracy: A Review and Implications for Developing Countries
Carsten Fink, Keith E. Maskus, and Yi Qian

Measuring Violent Conflict in Micro-level Surveys: Current Practices and Methodological Challenges
Tilman Brück, Patricia Justino, Philip Verwimp, Alexandra Avdeenko, and Andrew Tedesco

Promoting Women’s Economic Empowerment: What Works?
Mayra Buvinic and Rebecca Furst-Nichols

Employer Voices, Employer Demands, and Implications for Public Skills Development Policy Connecting the Labor and Education Sectors
Wendy V. Cunningham and Paula Villaseñor

Revisiting the “Cash versus Food” Debate: New Evidence for an Old Puzzle?
Ugo Gentilini

The Price Is Not Always Right: On the Impacts of Commodity Prices on Households (and Countries)
Daniel Lederman and Guido Porto
The Economic Effects of Counterfeiting and Piracy: A Review and Implications for Developing Countries

Carsten Fink, Keith E. Maskus, and Yi Qian

Policy makers around the world recognize the potentially harmful consequences of trademark counterfeiting and copyright piracy. Indeed, many countries have recently initiated policy reforms to strengthen the enforcement of intellectual property rights (IPR). Further, minimum standards of enforcement have been incorporated in many international treaties, especially trade agreements. This emphasis on enforcement raises basic questions about the actual impacts of IP rights infringement, which differ across the types of IPR and economic sectors. We review the academic literature and other studies in the public domain to evaluate what has been learned about these socioeconomic effects, with an emphasis on developing countries where possible. We also identify important gaps in our understanding of the consequences of counterfeiting and piracy and develop recommendations on how governments might collect data and conduct studies to better inform IPR enforcement policy.

Counterfeiting, Piracy, Intellectual property rights. JEL codes: L14, O31, O34

Intellectual property rights (IPR) cover a broad array of legal rules affording individuals and enterprises, or their licensees, exclusive opportunities to make, copy, and sell products and technologies arising from their inventive and creative work. These rules range from patents on new inventions to trademarks and geographical indications certifying the origin of goods. These rules also include copyright on creative and literary works, including software and digital transmissions. Because these rights are national in scope, how they are defined and protected varies considerably across countries for a variety of reasons (Maskus 2012).

Despite their broad similarity as a means of supporting markets for knowledge, various forms of IPR embody different economic and social objectives. It follows
that their infringement, or actions taken by third parties in violation of exclusive rights, interferes in varying degrees with attaining these objectives. Producing an imitative good that infringes on one’s patent rights may diminish dynamic incentives to invest in R&D, for example, though it can also bolster market competition in a static sense. Moreover, there is an important international element: if IPR infringement weakens the profitability of largely foreign enterprises, why would governments choose to invest in costly enforcement activities?

The particular forms of IPR infringement that have attracted the greatest global policy interest are counterfeiting, which is the illegal use of trademarks to sell unauthorized goods, and piracy, which is the duplication and sale of copyrighted goods without the copyright holder’s permission. Policy makers around the world recognize the potentially harmful consequences of trademark counterfeiting and copyright piracy. First, infringements of intellectual property (IP) rights undermine trust in rules-based systems that are at the core of modern market-based economies. Second, counterfeit products can confuse consumers and may be ineffective or even dangerous, as with adulterated medicines and foodstuffs or non-certified engine replacement parts. Third, extensive counterfeiting can significantly deter the emergence of new products and firms, and piracy can block the development of creative industries, especially in poor countries where entry conditions are already weak.

Over the past two decades, many governments—including those of developing countries—have initiated policy reforms to strengthen the enforcement of IPR (Maskus 2012). In addition, minimum standards of enforcement have also been incorporated in many international treaties, especially trade agreements, as IPR owners in wealthier economies push for stronger global protection. For example, the terms of the Trans-Pacific Partnership (TPP), while still under negotiation, embody extensive expectations for stronger border and domestic enforcement.

However, instituting a policy framework for combating counterfeiting and piracy poses substantial challenges. The economic effects of infringements differ across different IPR and economic sectors (Fink 2009), and it is important to understand these differences at the theoretical and empirical levels to set priorities for IPR enforcement policy. Governments are invariably resource-constrained, and completely eradicating violations of IP law—similar to violations of other types of law—is out of reach for even the best-resourced states. This is especially true for developing countries, where many public goods are underprovided and enforcement challenges exist in many areas of law.

Thus, in this article we review the academic literature and other studies in the public domain to evaluate what is known about the socioeconomic effects of counterfeiting and piracy. We also identify important gaps in the understanding of these effects and develop recommendations on how governments might collect data and conduct studies as inputs into the formulation of IPR enforcement policy. For the most part, we ignore differences and limitations in the effectiveness of IPR
enforcement measures, which are equally important but warrant a separate discussion. Where relevant, we draw inferences from these studies about important implications for economic development and policy making in developing countries. In particular, we highlight the importance of collecting better and more consistent data, conducting microeconomic studies of the impacts of infringements on local firms, and raising consumer awareness of the potential externalities they impose on domestic producers and the potential risks consumers face with sub-standard products. Governments may also wish to develop an impact evaluation framework to assess their strategic investments in enforcement activities. This could be particularly important in light of growing external pressures for stronger efforts in the context of regional trade agreements.

The paper is structured as follows. In the next section, we approach the topic from a theoretical perspective, first reviewing the different market failures underlying trademark and copyright protection, and then exploring what happens when these rights are violated. The discussion concludes with brief suggestions for future theoretical research. In the next section we review existing empirical evidence, including aggregate estimates of levels of counterfeiting and piracy, studies of the effects of IPR infringements in particular industries, and anecdotal evidence from selected case studies. In the final section we explore what the findings of the literature imply for policymaking in developing countries, and develop recommendations for data collection and the conduct of future studies on counterfeiting and piracy.

Theoretical Considerations

In general, intellectual property rights aim to remedy the failure of markets to provide an efficient allocation of resources. The relevant market failures that give rise to copyright and trademark protection vary. Accordingly, the effects of trademark counterfeiting and copyright piracy vary and warrant separate treatments.

Economics of Trademark Counterfeiting

The primary welfare effects of counterfeiting depend crucially on whether consumers are deceived into believing that a fake good is produced by the owner of the trademark. The likelihood of this deception depends on product characteristics and the nature of distribution channels. For example, consumers are usually unable to ascertain whether a pharmaceutical product contains the desired chemical ingredient, whereas most fashion shoppers can distinguish an original handbag from its fake clone. We discuss the implications of deceptive and non-deceptive counterfeiting in turn.
Deceptive Counterfeiting. The most relevant market failure giving rise to trademark protection arises from asymmetric information. For many goods, consumers do not have enough information to ascertain the quality of a potential purchase. A trademark guarantees that a product or a service originated with a particular producer. This is valuable information to consumers because it reduces both uncertainty and the costs of searching for particular quality levels. Indeed, producers compete by establishing a reputation for different levels of quality. In turn, trademarks serve as an indication of product quality, a crucial support for functioning markets. Thus, trademarks, geographical indications, and similar rights enable high-quality producers to distinguish themselves in the market, supporting investments in improved product or service quality.

Deceptive counterfeitters copy trademarks, logos, and designs to confuse consumers into believing they are buying the legitimate product. In the presence of information asymmetries, deceptive counterfeiting is virtually certain to reduce economic welfare (Grossman and Shapiro 1988a). Those who discover they bought a fake good realize less consumption value than the price they paid for it (Liu et al. 2005). Rational consumers, aware that fake goods are on the market but are indistinguishable from originals, are unwilling to pay the full price of a high-quality good. This problem undermines the incentive for producers to invest in higher quality and may destroy markets for high-quality goods (Akerlof 1970; Qian et al. 2013).

For developing countries, where deceptive counterfeiting sometimes reaches high levels, at least two significant costs arise (Maskus 2012). First, extensive anecdotal evidence suggests that the likelihood of infringement deters market entry by local enterprises, thus limiting the growth of small and medium-sized enterprises. In this context, counterfeiting can be anti-competitive in a dynamic sense, even in poor countries.

Second, for certain products, the consumption of deceptive fakes can generate negative externalities. Examples would be an increased risk of disease transmission in the case of pharmaceuticals, or environmental degradation in the case of industrial chemicals. Indeed, the distribution of counterfeit drugs poses a significant public health challenge in many developing countries, especially in Africa. The presence of negative externalities calls for public action against counterfeit products, independent of any private incentive for enforcing trademark rights.

Non-deceptive Counterfeiting. A more complex situation arises where consumers are perfectly able to ascertain the quality attributes of goods at the moment of purchase and therefore know whether they are buying a counterfeit. Because no information asymmetry prevails, trademarks, at first sight, seem irrelevant. However, for so-called status goods the display of the producer’s name confers prestige on the purchaser, yielding utility beyond that from the goods’ functional characteristics. Counterfeiting of status goods can reduce the prestige value of genuine versions.
Such value is real, as is manifested by the substantial marketing outlays of producers of luxury goods and the prevalence of counterfeits.

Grossman and Shapiro (1988b) analyze the effects of non-deceptive product counterfeiting. In their simplified model, status value depends negatively on the number of consumers who purchase a product bearing the same brand name, whether genuine or fake. In this setting, the social welfare consequences of product counterfeiting become theoretically ambiguous. Producers and consumers of the genuine product are made worse off by counterfeits because their presence reduces the status value, and therefore the price, of the genuine product. However, producers and consumers of the counterfeit product are better off. In particular, consumers of fake products gain status value without paying the full price of the genuine product. The overall effect on social welfare is an empirical question.

**Strategic Responses to Counterfeiting.** More recent analysis has identified further interesting features of markets with fake goods. For example, Qian (2014a) models the impacts of entry by counterfeiters on the subsequent responses of genuine producers. Because counterfeiters are usually an inferior imitation of the authentic product, the theory adopts a vertical differentiation setup. In particular, the theory incorporates two layers of asymmetric information potentially generated by counterfeiters: (1) counterfeiters fool buyers, and (2) buyers of counterfeit goods fool their peer consumers by pretending to have high-status goods. One key prediction is that entry by counterfeiters would induce a genuine producer to upgrade product quality and raise the product’s price if and only if the entrant’s quality is lower than a threshold level. In essence, genuine trademark owners must raise quality and price (and reduce sales) to attract demand in the market, which is consistent with evidence from China discussed in the next section. Note, however, that the possibility that original producers may invest in higher quality versions to compete with counterfeiters does not imply higher economic welfare, or that the induced innovation is efficient. Qian et al. (2013) further show that competition from counterfeiting induces branded firms to invest more in improving searchable quality attributes such as appearance, and less in enhancing experiential attributes, such as functionality.

Competition from counterfeit products can either lower the prices of genuine goods due to substitution in demand or raise their prices and production costs in order to move up-market. In addition, Qian (2014a) demonstrates how authentic producers could invest in “self-enforcement” strategies to limit the competition from counterfeit goods. Such strategies include conspicuous packaging of products and specially licensed company stores.3

**Longer-term Welfare Considerations.** Counterfeiting can encourage certain forms of innovation among genuine producers in an attempt to differentiate their products from imitations. This process is costly, however, perhaps especially in poor countries.
where the margin of high-income consumers seeking high-quality genuine goods is relatively small. More fundamentally, under both deceptive and non-deceptive counterfeiting, producers of genuine products realize lower profits. Facing this problem, original trademark owners become less willing to invest in developing new goods, ultimately reducing the rate of new product development and lowering both entry of new firms and consumer welfare (Aghion et al. 2005; Scherer 1967). We conclude that on balance, extensive counterfeiting is likely to harm consumers and impede growth in poor and emerging economies.

**Economics of Copyright Piracy**

The need for copyright protection comes from three characteristics associated with creative works such as books, music, films, and software. First, these products can be reproduced at low marginal cost; in the case of the Internet, reproduction occurs at virtually zero marginal cost. The second characteristic is that their consumption is non-rival; many people can enjoy them without diminishing availability for others. The third characteristic of such works is that the creative process may require costly sunk investments. If competitive market forces were to provide creative works at marginal cost, actors, artists, authors, producers, and publishers would have little financial incentive to undertake those investments. The copyright offers a solution to this problem by affording producers of creative works the exclusive rights to authorize and control their reproduction. In turn, genuine copies may be sold at a profit, allowing for the financing of creative investments.⁴

In light of this essential economic rationale for protection, what are the implications of copyright piracy? The basic answer is that high levels of copyright infringements undermine the incentives for creating and licensing new works, ultimately lowering social welfare. Unauthorized copying tends to flourish where originals command a high price, copying is cheap (as with digital goods and internet transmissions), and the likelihood of detection or meaningful punishment is low. In such circumstances, widespread copying can significantly diminish the incentives to develop and produce new creative and literary works. This barrier seems particularly relevant for indigenous music and software industries in developing countries with high rates of piracy, where copiers pervasively target local performers and emerging enterprises (Maskus 2000; Penna et al. 2004).

Beyond this basic framework a number of theoretical nuances should be mentioned. For example, theory suggests that piracy can raise or lower the price of original works (Besen and Kirby 1989; Takeyama 1997). If consumers who prefer to buy originals have sufficiently inelastic demand, it is possible for the prices they pay to be higher in the presence of pirated copies. The reason is that copying would induce original producers to concentrate their sales on this group, which pushes up
price. Overall sales of legitimate copies are likely to be lower, however, because copyright owners would choose not to sell to the remaining market segment.

An example is the model by Harbaugh and Khemka (2010), who point out that if copyright enforcement efforts are focused on larger users with higher valuations for software, such as public enterprises, universities, and government agencies, the original IP owners can charge higher-than-monopoly prices to those users. The remaining users in the market then purchase pirated copies. When enforcement is spread more broadly throughout the market, however, original producers will sell more legitimate copies and prices will fall, even as the volume of pirated goods diminishes. In this situation, it is possible to observe more genuine goods on the market and an increase in overall consumer benefits.5

Another reason prices may rise in the face of piracy is that unauthorized copying may induce copyright holders to add features and functionality to genuine copies to distinguish them in the marketplace. In turn, these features support higher prices, which consumers of originals are willing to pay, whereas pirated copies command lower prices (Varian 2000). An early variant of this concept is detailed by Liebowitz (1985), who presents empirical evidence that academic journals raised their prices after the introduction of photocopying because they had become more valuable to consumers. A similar process is underway in academic publishing in the digital age, with dramatic increases in subscription prices for paper copies (Maskus 2012).

Finally, note that books, journals, music, and videos are “information goods” that can be shared across multiple uses through copying, rentals, and literal sharing. This fact affects strategies of the content producers and copyright owners, who must decide the formats, timing of releases, functionality, and prices of such goods. According to Varian (2000), copyright owners are likely to sell fewer copies at higher prices to manage the sharing problem. This strategy is likely to be profitable when there are low transaction costs in sharing (e.g., movie rentals) and when sharing markets permits separation of higher-valuation and lower-valuation users. The former will receive goods faster at higher prices, and the latter will have delayed access at lower prices or will enter lower-quality and, perhaps, pirated markets.

**Network Effects.** Piracy takes on interesting features in the presence of goods with network externalities, in which one consumer’s valuation of a creative work increases with the number of other consumers owning the same product. Computer programs are an example, as the ability to share files with other users increases the attractiveness of a particular software package.

What happens if creative works subject to network externalities are pirated? Suppose that original and pirated copies are perfect substitutes, or at least highly interoperable as in the case of software. Consumers of original copies gain because the presence of pirated copies enlarges the network with which they can interact. In addition, consumers of pirated copies will benefit from inexpensive access to the work.
The effects of copyright piracy on the producer of the creative work and overall welfare are ambiguous. As Conner and Rumelt (1991) and Takeyama (1994) show, it is theoretically possible for the copyright owner to reap higher profits from certain levels of piracy. The intuition is that the enhanced network value of the product may allow producers to charge a higher price for originals. Here, copyrights support profitable price discrimination. In theory, the copyright owner could achieve the same outcome by giving away a certain number of original works. In practice, this would not work because all consumers would expect a free legitimate copy. However, consumers differ in their willingness to use pirated copies, as determined, for example, by income levels. This difference allows the producer to segment the market and choose the profit-maximizing combination of price and network size.6

Effects on Tax Revenue and Employment

In policy discussions on trademark counterfeiting and copyright piracy, effects on tax revenue and employment have gained in importance, especially in richer countries where firms see their IPR infringed on abroad. From an economic perspective, these effects can be evaluated either in a shorter-term partial-equilibrium setting or a longer-term general-equilibrium setting.

The overall partial-equilibrium employment effects are theoretically straightforward and depend on how output shifts between genuine and illicit producers and the labor intensities of original and pirated copies. The same holds for tax revenues, which are nearly certain to fall because sales of pirated and counterfeit products occur in informal markets where taxes and import tariffs are usually not collected, and because sales of legitimate producers likely decline. No existing theoretical research examines the tradeoffs between employment in the formal and illicit informal sectors, including the setting of wages, the existence and impact of social safety nets, and the nature and length of possible unemployment spells.

The long-term general-equilibrium effects of IPR infringement are often ignored in policy discussions, but understanding them better is crucial. Again, direct impacts in affected sectors are surely negative. However, workers losing employment likely find other jobs, and governments facing a revenue shortfall likely adjust their tax structure to finance public spending. The key question is how workers and the efficiency of the tax system fare in the counterfactual equilibrium. No formal modeling work seems to exist in the literature that would offer conceptual guidance.

Impacts on Trade, Foreign Direct Investment, and Licensing

The theoretical literature on the relationships between IPR protection and international trade, FDI, and licensing is deep and sophisticated.7 These linkages
are important because they govern how policy reforms and enforcement affect market-based inward technology transfer (ITT) into developing economies. However, economic theory in this area focuses almost exclusively on IPR as technology protection, for example patents and trade secrets, and has ignored trademark counterfeiting and copyright piracy. Further, no models have considered the role of stronger enforcement as being distinct from basic legal reforms.

Thus, a comprehensive review lies beyond the scope of this paper and we simply state the primary theoretical findings. Because multiple channels exist for transferring technology, IPR in the recipient country has both scale effects and composition effects. Regarding scale, multinational enterprises (MNEs) seeking to deploy a new technology or sell a new good face a basic tradeoff when local IPR is enhanced (Maskus and Penumarti 1995). Because local firms face higher imitation costs, their outputs are reduced and MNEs sell more, whether through trade or FDI. At the same time, MNEs face less elastic demand from consumers who continue to buy their goods, thus permitting higher prices at reduced (monopolistic) volumes. This tradeoff between market expansion and market power is fundamental to the theory of intellectual property and ITT. This tradeoff extends to licensing as well: MNEs can expand the number of contracts for technology use and production rights but they can also raise royalty rates (Branstetter et al. 2006). The key point is that IPR reforms have ambiguous impacts on ITT, at least in theory, making it an empirical question.

As for the composition of ITT, economic theory suggests that MNEs shift their emphasis from exports at weak IPR levels to FDI and then licensing as protection improves (Markusen 2001; Javorcik 2004). This theory assumes (with little evidence) that imported high-technology goods are relatively difficult to imitate and thus carry their own protection. As IPR is strengthened, MNEs gain more confidence that local affiliate production cannot readily be copied, and become more willing to take advantage of cost reductions through FDI. However, technology remains internal to the firm for purposes of control. At yet higher protection levels, MNEs become willing to license their technologies to local partners, thereby externalizing their knowledge assets. A related factor is that stronger patents and trademarks increase the confidence that contracts are enforceable, making arm’s-length licensing more attractive.

One additional strategic response (among many) is worth mentioning here. Specifically, where MNEs fear losing their technological advantages to imitation and their trademark reputations to counterfeiting, they may fragment their production among multiple local plants at various locations to diminish information flows among them and gain more ability to deter infringement. This approach, which has been common among MNEs operating in China (Maskus 2000), is inefficient and likely reduces net technology transfer. For example, Sun et al. (2010) demonstrate theoretically that an MNE worried about imitation in an emerging market may limit the range of technologies it offers to deter entry by local imitators.
As mentioned, there are no theoretical studies focusing on FDI, licensing, and ITT with respect to counterfeiting and piracy, our focus here. However, many of the tradeoffs above surely apply. For example, reducing trademark infringement can lower the local selling costs of MNEs and expand their markets, whether through trade or FDI. Regional fragmentation of production is a costly reaction to counterfeiting, as is the need to engage in private enforcement. And, because of weak copyrights, MNEs may deploy expensive software to deter copying or pay more for programs on the market, as noted above. Presumably, these factors are deterrents that companies take into account when deciding on investment locations.

It is important to note that high-technology imports and FDI support localized learning and productivity gains in developing countries (Keller 2004; Branstetter et al. 2006), which are important for supporting productivity growth. To the extent that weak IPR diminishes such prospects, they have a negative dynamic impact on development.

Recommendations for Future Theoretical Research

While the economic literature offers useful theoretical guidance on the economic effects of counterfeiting and piracy, several issues deserve further exploration. We highlight three such issues here. First, as we described previously, the presence of (non-deceptive) counterfeit goods exerts an externality on consumers of originals. In their analysis, Grossman and Shapiro (1988b) assume that this externality is negative and takes the form of counterfeits diluting the exclusivity value of status goods. This assumption is appropriate for some classes of goods but not for others, and survey evidence suggests there is substantial heterogeneity in consumers’ decisions to purchase counterfeit versus authentic products Wilcox et al. (2009). Other types of externalities are conceivable, and it would be important to analyze how their presence affects consumers, the strategic reactions of producers, and overall welfare.

Second, hardly any theoretical study has formally analyzed the short- and long-term employment effects of counterfeiting and piracy. It would be especially important to obtain conceptual guidance on how the different nature of formal and informal labor markets affects the welfare of different types of workers. Such guidance would seem especially important for developing countries, where informal employment often accounts for a larger share of total employment. Developing a better understanding of broader labor-market linkages may also be important for designing effective IP enforcement policies Fink (2009).

Finally, it would be interesting to focus theoretical attention specifically on counterfeiting and piracy in terms of their potential impacts on technology transfer and localized learning. Little is known, for example, about how the enforceability of IPR affects relationships between MNEs and local suppliers and distributors. Nor is there
much analysis of how piracy and counterfeiting influence global franchising decisions, which often embody considerable knowhow regarding market organization.

**Review of Empirical Evidence**

Economic theory is helpful for rigorously examining the likely effects of trademark counterfeiting and copyright piracy. However, as we noted previously, some effects are theoretically ambiguous and require empirical investigation. In addition, even when the overall direction of effects is conceptually clear, empirical insights into the scale of effects are needed for policy makers to set priorities for public policy.

Notwithstanding the need, generating credible empirical evidence poses a significant challenge. Counterfeiting and piracy, by their very nature, are illegal activities and thus largely escape official statistical recording. The basis for empirical research is thus weak, and studies must rely on indirect official data, selected information supplied by rights holders, and original surveys to gather evidence. In addition, welfare impacts such as producer and consumer surplus gains or losses are difficult to measure at both the individual and the aggregate societal level. Finally, the limited empirical evidence that is available comes mostly from developed countries. While this evidence can still inform policy makers in developing countries, there are special characteristics of poorer economies that caution against the direct applicability of this evidence—notably, different parameters of consumer demand and different legal and institutional frameworks.

In what follows we critically review relevant empirical studies, and focus on the extent to which the underlying methodologies produce credible results. We begin with a review of aggregate estimates of the level of counterfeiting and piracy. We then examine studies that have evaluated the welfare impact of these activities at the micro level, and review the studies that have specifically focused on the effects of Internet file sharing. We conclude by evaluating the advantages and disadvantages of alternative empirical approaches.

**Aggregate Estimates of Levels of Counterfeiting and Piracy**

We begin the review of aggregate estimates by considering efforts to ascertain the incidence of piracy and counterfeiting in international trade. Focusing on trade seems promising, because goods undergo processing when they cross borders. In particular, customs authorities investigate suspected infringements of IP rights, thereby generating indirect data on the incidence of pirated and counterfeit goods in different product categories and for different exporting and importing economies.

The most comprehensive study attempting to quantify the importance of counterfeit and pirated goods in international trade was conducted by the OECD (2008).

Fink et al. 11
This study estimated the value of pirated and counterfeit goods in international trade at U.S. $200 billion, or approximately 2% of global merchandise trade in 2005. In an update of this study, the OECD (2009) estimated the value of illicit goods in international trade as having grown to $250 billion in 2007 (largely reflecting the growth in worldwide trade). These estimates do not include the value of Internet-related piracy of digital goods, which easily cross borders but are almost impossible to measure and, in any case, cannot be counted in any merchandise trade categories.

The OECD (2008) report recognizes that its estimates are only “a crude indicator” of the value of counterfeit and pirated products in international trade. This caution is warranted. For instance, the OECD estimate employs data on interceptions and seizures by customs authorities in selected countries to estimate the proportion of counterfeit and pirated goods in different product groups and across exporting economies. However, these values only provide information about the relative incidence of pirated and counterfeit goods across product groups or across exporting economies. To arrive at their estimates, OECD staff use a value of the absolute incidence of counterfeit and pirated goods in one particular “fix-point” product group (e.g., apparel, leather articles, tobacco products). The value in question is not based on any hard data, but rather seems to reflect the best guess of OECD staff.8

Notwithstanding these methodological caveats, the evidence on the relative incidence of counterfeit and pirated goods across product groups is insightful. It shows, for example, that trade in illicit goods is concentrated in a small number of “sensitive” product categories; the top five product groups (at the two-digit Harmonized System level) account for more than three-quarters of all customs seizures.

Industry Studies. Moving beyond trade, a large number of studies—often sponsored by affected industries—attempt to quantify the incidence of pirated and counterfeit goods at the level of specific sectors and to estimate their effects on certain economic performance variables. The methodological approaches and quality of the data in these studies vary substantially. A detailed review of each of these studies would go beyond the scope of this article, but it is worth highlighting several common issues.

First, given the inherent difficulty of accurately measuring the prevailing levels of counterfeiting and piracy, the reliability of any underlying estimate is in many cases questionable. This difficulty applies to the results of original survey work. For example, industry associations in the major copyright-intensive industries publish annual surveys of piracy rates in major countries around the world. These questionnaires may collect useful information about attitudes toward piracy and means of unauthorized copying, but the resulting estimates of illegitimate use need to be treated with caution.

As a specific example, the Business Software Alliance (BSA) released their ninth study in May 2012, in which they argued that the global piracy rate for PC software
stood at around 42% in 2011, representing more than $60 billion (BSA 2012). This finding came from a study of more than 110 countries, with piracy rates ranging from approximately 19% (United States and Japan) to over 90% (Georgia and Zimbabwe). To compute these rates, the study relied on surveys of sampled consumers and businesses in a smaller set of 28 countries to determine the number of computers in use and the share of legitimately procured software. The study then estimated piracy rates for countries not in the sample surveys by means of a correlation between software usage and an “information development index” published by the International Telecommunications Union. The extent of bias in the estimates for these non-sampled countries arising from both the underlying approach and the application of the index is unknown. Regarding the latter, considerable variation is likely to exist in its applicability to countries of differing economic and social characteristics, as is the case for most such “development” indexes.

Second, estimates of the economic effects of counterfeiting and piracy often do not account for demand responses—the possibility that not all consumers would switch to the original product if the counterfeit or pirated version were unavailable. Again, the BSA (2010) piracy estimates illustrate this point. Based on estimates of the volume of unlicensed software, the BSA computes its commercial value by applying a vector of prices for various kinds of computer programs including retail, licensed, and open-source. This total value is then taken as a measure of lost sales for legitimate software producers in each country. This approach assumes that each unit of pirated software would be fully replaced by a purchased version if copying were eliminated. That the BSA’s estimates of volumes and values foregone are likely substantially overstated emerges from a recent study estimating demand responses for one popular software product, Microsoft Office. Relying on original conjoint survey data from college students in Hong Kong, Leung (2013) estimated a discrete choice demand system for Microsoft Office from both legal and illegal sources; he concludes that the true gain from eliminating all sources of piracy is only 15% of the BSA’s estimated cost of piracy.

Third, the counterfactual market equilibrium analyzed is often not clearly spelled out, nor does it seem realistic. In particular, studies provide estimates of the economic implications of certain levels of counterfeiting and piracy, and associate those levels with employment and tax losses without considering the longer-term general equilibrium responses outlined previously. Possible employment effects in the informal sector are almost always ignored.

Some studies attempt to consider cross-economy impacts, at least in an input-output framework. Consider, for example, a widely-cited calculation of U.S. copyright industry sales, employment, and tax revenue losses (see Siwek 2007). The author employed industry-generated measures, such as those from BSA and LEK, of global sales losses in motion pictures, recorded music, software, and video games. Siwek considered the sum of these losses to be an underestimate because not all
countries are covered in all industry estimates. Making some adjustments for substitution between pirated and legitimate copies, and valuing all lost sales at legitimate prices, the study suggested a direct total loss of $25.6 billion for U.S. industries (including retailers).

Siwek (2007) then applied “multipliers” available from the U.S. Department of Commerce, derived from its input-output model of the U.S. economy, to determine how this lost sales figure translates into total (direct plus indirect) losses in economic activity and total employment lost. These basic calculations resulted in a total sales loss for the U.S. economy of $58 billion per year, along with 373,000 jobs lost as a result of global piracy. All of that translated further into an earnings loss of $16.3 billion and reduced tax collections of $2.6 billion. These are large figures but must be kept in context.

Although the multipliers account for cross-industry impacts in a static sense, the basic input-output approach again does not consider general-equilibrium linkages in the economy. If there really were $25.6 billion in reduced demand for U.S. copyright goods, consumers would focus their expenditures elsewhere in the domestic and international economies, while workers would move to alternative employment and output and tax payments would rise there. Thus, Siwek’s (2007) study offers an assessment of the gross losses in sales and employment but greatly overstates the net losses. As we described previously, analysts would need to develop a general-equilibrium model for the latter calculations with endogenous changes in sectoral demands, output, and trade, while computing earnings losses at the difference between with-piracy and without-piracy wage rates.

The United States International Trade Commission Analysis. Because of the methodological shortcomings outlined above, we are skeptical that the results of the studies reviewed to this point offer useful guidance to policy makers. A more careful study was performed by the United States International Trade Commission (USITC: 2011); this project focused exclusively on the effects of IPR infringement in China, where concerns among major technology, fashion and content creators in the United States, the EU, and other rich countries are primarily focused.

The USITC’s figures on gross sales lost due to Chinese infringement were taken largely from surveys of American firms. These firms claimed sizable damages, largely due to copyright piracy and trademark counterfeiting, though issues of indigenous innovation and industrial espionage were also considered. Analysts at the USITC applied these reported reductions in demand to their well-regarded computational general-equilibrium model. Such models are more flexible and more informative than simple computations with input-output tables, for they permit endogenous responses in prices, outputs, factor use, trade, and product demands. At any rate, the USITC model computed losses of nearly one million jobs in the United States in the sectors most affected by Chinese infringement. There was,
however, considerable uncertainty about this estimate depending on the scenarios considered.

After having conducted this review, we observe that capturing the impacts of global infringement of IPR is a complex and uncertain problem that could use far more analytical and data attention. Moreover, there are no reliable computational general-equilibrium studies on how infringement affects the economies of developing countries. This is not to say that the economy-wide effects are negligible and do not deserve the attention of policy makers in both developed and developing countries. We merely emphasize the challenge those policy makers face in addressing counterfeiting and piracy problems with limited empirical guidance on offer.

**Micro-data Studies on the Effects of Counterfeiting and Piracy**

Although obtaining aggregate estimates of the levels of counterfeiting and piracy is important for understanding the scale of the problem, effective policy cannot be devised without considering the micro level of firms and consumers.

**Marketing Analyses and Consumer Surveys.** Several marketing studies have been conducted to help understand consumer psychology and behavior that characterize the demand-side origin of counterfeiting. Although many of these are surveys or case studies, a few are experiments. Generally, experiments—whether in the field or in the lab—generate more credible evidence, but they also face methodological limitations. Field experiments that randomly assign people to control and treatment groups potentially have higher external validity because participants are real-world consumers making statements or decisions about the products of interest in (close to) real-world settings. However, compared with a lab experiment, in a field study it is more difficult to avoid interferences among participants and to exert tight control over the randomized treatment manipulations. Field experiments therefore generate results that are more susceptible to these “internal validity” concerns.

As an early study in this literature strand, Bloch et al. (1993) examined the consumer’s role as an “accomplice” in the proliferation of product counterfeiting. These authors presented results from a field experiment showing that a surprisingly large proportion of adult consumers will select a counterfeit garment over the genuine good when there is a price advantage. The researchers also found that a significant price gap and negative attitudes toward large branded companies were the main factors driving counterfeit demand.

In more recent work, Wilcox et al. (2009) conducted lab experiments and found that products with branded logos were more likely to be valued for their ability to help consumers gain social approval and status. In turn, this characteristic makes consumers more willing to buy such products from counterfeitors.

---

*Fink et al.* 15
Looking further into the determinants of why consumers purchase non-deceptive counterfeit products, Vida (2007) reports on a consumer survey conducted in Slovenia. The results suggested that personal religiosity and education tend to reduce one’s willingness to buy fake goods. Such was not the case in Kwong et al. (2003), who studied consumer attitudes toward pirated CDs in Hong Kong. These researchers found that approximately 70% of the respondents stated that they bought pirated CDs and that they thought this generated high social benefit and low social cost. However, both studies concluded that gender matters, in that men are more willing to buy counterfeits than women. Bian and Veloutsou (2008) surveyed consumers in China and the United Kingdom, and found that perceived financial, physical, and performance risks from consuming counterfeits were moderate. Still, consumer willingness to buy fakes varied with each product and each country, which makes it difficult to draw general conclusions.

Finally, Casola et al. (2008) conducted three studies of consumer behavior in hypothetical situations. All three studies revealed that consumers were less willing to buy counterfeits when the victim was an individual rather than an organization. The average consumer would buy fakes if they cost about one-third of the genuine price. Finally, consumers’ willingness to buy counterfeit goods decreased if they were informed of the harmful impacts imposed on others. Thus, spreading awareness about the negative consequences of counterfeiting may be effective at deterring demand.

Field Studies. Overall, studying consumer psychology is perhaps more useful for understanding why there is a demand for counterfeits than for estimating the economic impacts of counterfeiting. For the latter, it is more fruitful to collect field data over time to generate a sample panel on actual branded firms and their infringers. This is a challenging task because of the illicit nature of counterfeiting and the confidentiality concerns of both genuine and infringing companies. Nonetheless, a series of recent studies have been carried out in this direction.

Qian (2008) offers the most comprehensive empirical investigation on the economic impacts of counterfeits. This study also stands out in that it is one of a relatively few studies that focus on a developing country. Qian collected original panel data from Chinese shoe companies from 1993 to 2004 (after government enforcement efforts were reduced in the footwear sector) through stratified random sampling to identify and measure the effects of competition from counterfeit goods on the prices, qualities, and other market outcomes of authentic footwear. The results show that brands experiencing less government enforcement action differentiated their products by introducing higher visible quality. These brands also engaged in forms of self-enforcement by establishing vertical relationships with licensed downstream retailers (“company stores”) to control their sales, and setting higher prices.
to signal their legitimacy in the marketplace. All these reactions helped reduce counterfeit sales in the period studied.

It should be clearly noted that all the innovation and differentiation strategies mentioned above are costly. Although the authentic prices dropped initially under competitive pressure from counterfeits, the long-term effects on driving up authentic-product prices kicked in at various times for different firms. Larger firms with more human capital and research-and-development resources were faster in responding and differentiating their products from counterfeits. Firms with more exports responded slower, perhaps because they were more diversified and less influenced by domestic counterfeits (Qian and Xie 2010).

Field panel data of the kind used in the Chinese footwear study provide direct measures of how various economic variables change over time within each brand, but there are limitations that mean the results should be treated with caution. For example, even when field panel data are available for estimating how economic performance changes in response to counterfeiting, the underlying mechanism may be difficult to identify. Qian (2014b) explored the sales impacts of counterfeiting based on a combination of field data and lab experiments, and found that counterfeits have both a positive advertising effect (inducing more awareness of the brand) and a negative substitution effect on authentic products (reducing their sales). The advertising effect dominates the substitution effect for high-end authentic product sales, and the substitution effect outweighs the advertising effect for low-end product sales. The advertising effect is found to be larger for fashion products and for brands that were less well-known at the time of infringement. Another important observation about consumer attitudes toward counterfeiting is that a high degree of income inequality seems to cause a greater demand for fake goods (Qian and Rucker 2013).

Finally, several micro-data studies examine the economic effects of copyright piracy. Notably, Hui and Png (2003) estimated the effects of physical CD piracy on the legitimate demand for recorded music in an econometric setting. These authors’ model expressly accounts for the demand linkages outlined in the previous section (sales of pirated copies could stimulate more demand for originals). Their econometric results showed that piracy had a negative net effect on the demand for legitimate music, though Hui and Png’s estimate of foregone sales by copyright holders was 58% lower than the music industry’s estimate (see International Federation of the Phonographic Industry 2003). The latter assumed a one-to-one substitution of pirated with legitimate purchases, whereas Hui and Png’s model allowed for a share of price-sensitive consumers to choose not to switch to legitimate goods in the counterfactual scenario. The Hui and Png analysis is limited, however, by an assumption that copyright holders do not adjust prices in response to lower piracy.

That demand is responsive to prices charged for pirated and original copies also emerges from an experimental study of the music consumption behavior of...
university students. Maffioletti and Ramello (2004) found that in general, students’ willingness to pay for a CD is lower than the market price of a legitimate copy. Thus, the hypothetical elimination of piracy would not expand sales of legitimate copies on a one-to-one basis at existing prices. At the same time, the study reveals that students’ willingness to pay for a pirated CD was substantially greater than its marginal cost. This finding clearly points to the possibility that copyright holders might respond to stronger copyright enforcement by lowering their prices to capture a larger number of consumers.13

The Effects of File Sharing on the Internet

One of the most frequently studied aspects of copyright infringement is how unauthorized downloading and file sharing have affected sales of recorded music offered by major music publishers. It should be noted that this “end-user piracy” is sometimes different in intent from commercial piracy. In particular, many people who download music for free and put music files on their computers in forms that can be readily shared by others rarely attempt to make money from these actions.

Furthermore, some surveys suggest that many such users see little wrong with freely taking and sharing digital information products. For example, a recent survey of 1,607 people in the United Kingdom found that more than 80% admitted to having downloaded at least one file without authorization, and 47% said they did not think it was a crime (Bonnett 2010). In a survey of 1,000 U.S. college students in 2003, 69% said they had downloaded music, and 75% of those indicated they never paid for it (IPSOS Public Affairs 2003). Moreover, 76% responded that they would download even if they suspected the music file was illegal. These perceptions are among the reasons this form of unauthorized use is exceptionally difficult to control by copyright owners.

Regardless of the motivation, downloading and file sharing through peer-to-peer (P2P) networks is common. The OECD (2005) calculated that one-third of Internet users in its member nations have downloaded files from P2P networks. Regarding the effects of downloading and file sharing, U.S. shipments of recorded music on CDs fell by 25% between 2000 and 2005 (Recording Industry Association of America 2006), and representatives of the music publishers blamed this trend on illegal downloads and P2P file sharing. Most recently, a global music industry association noted a 7% drop in music sales in 2009, though it was unclear how much was due to unauthorized activity (International Federation of the Phonographic Industry 2010). That report claims that P2P piracy accounted for more than 20% of global Internet traffic, with higher shares in Latin America and Europe.

This proliferation of file sharing has negatively affected authorized music sales. For example, Rob and Waldløgel (2006) collected data on albums obtained through downloading or purchase and used surveys among students at the University of

18

The World Bank Research Observer, vol. 31, no. 1 (February 2016)
Pennsylvania to measure their valuation of this music. Using changes in Internet
access as an instrumental variable, they found that downloading reduced the per-
capita student expenditure on hit albums from 1999 to 2003 from $126 to $100,
though it raised per-student welfare by $70.

The most widely debated analysis surrounds the work of Oberholzer-Gee and
Strumpf (2007), who suggest that downloading could, in theory, increase or
decrease album sales due to the peering effects we outlined previously. These
authors regressed the volume of sales by album on the number of downloads of
songs on the album and other factors. To control for a possible simultaneity bias,
they used as a primary instrumental variable the number of German secondary
school children on holiday when they had more time to wait for downloads. Their
results were striking; in a basic ordinary least squares regression, the coefficient of
sales on downloads was 1.09 and highly significant, suggesting a large positive
impact. However, after adding album fixed effects and first-stage instrumentation,
they could find no evidence that the number of times an album was downloaded
had a statistical impact on music sales, while the implied economic effects were
small.

This finding has been criticized on several grounds, primarily by Liebowitz
(2005; 2006; 2007). In their more recent study, Oberholzer-Gee and Strumpf
(2010) acknowledged that file sharing has reduced music sales, but by more
modest amounts than generally supposed. Furthermore, they argue that offsetting
increases in demand for complementary music services have increased incentives
for the creation of new music, books, and movies. In other words, file sharing may
be damaging to music publishers but not necessarily to artists themselves, espe-
cially those who are younger and less well-known. However, this possibility deserves
further study. Unfortunately, empirical studies on the effects of copyright piracy—
whether physical or online—are almost exclusively confined to developed countries.
To a good extent, this reflects the availability of data and the “home bias” of re-
searchers in developed economies. However, there are reasons to believe that copy-
right piracy is more widespread in poorer countries. For example, across a wide
range of countries, Fink (2009) reported a correlation coefficient between software
piracy rates and per capita GDP in 2004 of −0.89. This likely reflects poorer
consumers turning more frequently to pirated goods, the greater importance of the
informal economy as a source of employment, and less developed and more
resource-constrained legal systems.

Copyright-based industries exist in most developing countries. In some—such as
the Philippines, Malaysia, and Mexico—these industries account for larger employ-
ment shares than in developed economies. Copyright piracy is bound to affect the
prospects of earning a livelihood from creative and artistic activities in poorer
countries. One study on the music industry in West Africa reported anecdotal evidence
confirming both the high incidence of piracy and its negative effect on musicians’

Fink et al. 19
incomes (Penna et al 2004). However, this study also pointed to a large number of other barriers—from ineffective collection societies to tax policies—that render the recording activities unprofitable for most musicians. How a reduction in piracy levels alone would affect artists and those involved in the distribution of creative works remains an open question.

Conclusions

We conclude this review with some observations on what the literature suggests about the importance of policy initiatives for dealing with counterfeiting and piracy in developing nations. We then discuss the need for further work to improve data collection and analysis.

Implications for Policymaking in Developing Countries. As we have noted, most of the studies we reviewed above offer evidence about the impacts of counterfeiting and piracy in more developed economies, with the notable exception of Qian (2008; 2014b). Thus, readers may question their applicability to developing countries and policy authorities. Indeed, there are important differences to note. First, the classic “intellectual property industries” (high-end fashion, digital content, publishing, software, and research-based pharmaceuticals, food products, and plant varieties) play a smaller role in the production structures of poor countries. Rather, local firms may be more focused on imitation and copying, while consumers may be more open to buying fake goods. In this context, the immediate gains from stronger IPR enforcement are likely to accrue to foreign content providers and multinational enterprises, most of which develop the bulk of their intellectual property in rich economies. It is thus not surprising that many developing countries face demands for more stringent IP enforcement in bilateral and regional free trade agreements they negotiate with developed countries.¹⁵ Notwithstanding the broader welfare gains that such agreements could offer, policymakers in developing countries reasonably wonder whether it makes sense to devote scarce development resources to IPR enforcement.

However, it is important to point out that there is considerable and growing heterogeneity among developing countries. Rapidly emerging economies such as Brazil, Mexico, China, and Turkey see increasing interest among growing middle classes for higher quality goods protected by recognizable brands. There is also a rapid trend among innovative local firms toward brand development, investments in product variety and quality, and R&D expenditures in numerous industries, including software and digital content (Maskus 2012). In this regard, domestic interests in stronger IPR protection and enforcement are rapidly emerging and challenging the political-economy equilibrium that favors weaker enforcement. These countries presumably can learn from the experience of developed countries.
Second, even poorer countries have significant interests in stronger enforcement for at least two reasons. One is that their citizens disproportionately suffer from adulterated counterfeit medicines and ineffective or dangerous foods and beverages. Basic calculations suggest that this can be a major problem for sustaining public health in poor countries and that there is a positive payoff from working to reduce such risks (Maskus 2007). Another reason is that weak trademark enforcement can sharply diminish prospects for effective brand development, firm entry, and national marketing success among potentially innovative companies in poor countries. Indeed, local counterfeiting often targets domestic enterprises at least as much as foreign brands (Mertha 2005). Similarly, extensive unauthorized copying of music, designs, and other creative activities limits opportunities for artists and artisans to build local markets, ultimately harming their ability to sell globally as well. Thus, public investments in enforcement and creative infrastructures, perhaps in partnership with private entities, could help remove such barriers to entry. As a strategic matter, invariably scarce enforcement resources could focus on reducing deceptive counterfeiting (to deal with dangerous fakes) and educational campaigns on expanding consumer awareness of how unauthorized copying negatively affects local artists.

Recommendations for Data Collection and Future Empirical Work. Policy makers in countries at different stages of development are interested in better understanding the economic effects of trademark counterfeiting and copyright piracy. So what direction should future analytical work take? As discussed previously, many studies exist that aim to quantify the extent of counterfeiting or piracy at the aggregate level, be it sectoral or economy-wide. Most of these studies have serious shortcomings, both in the data employed and in the adopted methodologies. More important, aggregate estimates of the incidence of piracy and counterfeiting by themselves offer policy makers little guidance on appropriate IP enforcement policies.

More useful guidance emerges from micro studies that take into account the specific characteristics of different industries and comprehensively assess the effects of illicit activities on prices, consumer demand, company revenues, innovative behavior, and overall welfare. Such studies are still scarce. Finding reliable and representative data is challenging and creative approaches are needed to establish causality and derive credible estimates. Experimental studies offer one way forward, particularly in combination with broader empirical investigations. For example, researchers could gather field panel data to analyze the overall impacts of counterfeiting, and supplement them with lab experiments to understand the underlying mechanisms (see Qian 2014a; Qian and Rucker 2013).

The quality of any future empirical work will largely depend on the quality of the underlying data, however. Given the illegal nature of counterfeiting and piracy, collecting representative and systematic data on all relevant economic variables will
continue to be a challenge, and progress will only be gradual. Nonetheless, several promising avenues exist.

For example, more could be done to improve and harmonize data collection in the course of IP enforcement activities. Customs statistics are an obvious place to start because they are generally the most direct source of information about the extent of piracy and counterfeiting flows that cross international borders. Many customs offices already compile and publish data on seizures of counterfeit and pirated goods. Although the World Customs Organization has developed international guidelines for reporting on seized products, actual reporting practices differ substantially across countries. For example, some countries report the number of seizures, whereas others focus on the quantity of seized goods, or the number of consignments. Substantial differences also exist regarding the valuation of seized goods (World Customs Organization 2009).

From an analytical point of view, it would be important to achieve as much consistency as possible in the treatment of firms, products, and sectors. Data collection needs to be done consistently over time to track trends and analyze the impacts of changes in enforcement policy. Sampling interceptions data could be extended to detailed subcategories of goods to gain a better understanding of the scope of counterfeit trade.

Beyond customs statistics, there is scope for collecting more comprehensive data associated with domestic law enforcement activities, notably through the judicial system. To our knowledge, no systemic effort exists at the international level to promote and harmonize the collection of such data. Intergovernmental organizations can play a more active role in coordinating data collection initiatives, promoting the harmonization of reporting standards, and offering technical assistance to developing country governments.

At the domestic level, governments may want to consider adopting an impact evaluation framework, especially when new IP enforcement policies and measures are introduced. Continuous data gathering would be an essential component of such a framework. Finding appropriate control groups against which enforcement activities could be evaluated may prove challenging but is possible in certain circumstances—especially when new initiatives are first implemented on a pilot basis.

In addition, there is scope for conducting original surveys to support targeted analytical work. Surveys of rights holders on key performance characteristics as they relate to counterfeiting activity could usefully complement firm-level data available through industry or official sources. Depending on the industry, rights holders may also possess valuable information on the nature of, and markets for, counterfeit and pirated goods that was gathered in the course of their own market intelligence and enforcement activities. Surveys of rights holders could take place within existing cooperation programs between the public and private sectors. Such firm surveys should be designed to collect as detailed and comprehensive data as possible while
guaranteeing the anonymity of respondents when submitting sensitive business information.\textsuperscript{16}

In the longer term, as information from these surveys—especially across countries—becomes more standardized and comprehensive, econometric models could be developed to estimate important market parameters about which limited information exists. Most prominently, the demand characteristics and substitution parameters between legitimate goods and counterfeit goods are important to estimate to gain an accurate picture of damages imposed on rights holders and the associated effects on employment and research-and-development spending. Such estimates need to be subjected to significant sensitivity and robustness analysis to ensure that they can be deployed with confidence.

After these parameters are estimated, it would be feasible to develop computational general equilibrium models to assess the country-specific and international impacts of counterfeiting and piracy, particularly accounting for cross-industry effects and employment channels. Models of this sort would be particularly useful for assessing the potential impacts of additional policy efforts to counteract illicit trade and production. It is important to emphasize, however, that such models necessarily rely on strong assumptions that could be misleading in this context. For example, employment in the counterfeiting and piracy sector is largely informal, and few models can manage the linkages between informal sector and formal sector employment. Still, with appropriate caveats, these models should be an improvement over existing estimates of economy-wide losses in employment and tax revenues because those estimates do not account for aggregate relationships in the economy.

Notes

Carsten Fink is the Chief Economist, World Intellectual Property Organization: carsten.fink@wipo.int. Keith Maskus is a Professor of Economics, University of Colorado Boulder. Yi Qian is an Associate Professor of Marketing and Behavioral Science, Sauder School of Business, University of British Columbia. The authors contributed equally and are listed in alphabetical order. Helpful comments from Intan Hamdan-Livramento, Sacha Wunsch-Vincent, and two anonymous referees are gratefully acknowledged. The opinions expressed herein are those of the authors and should not be attributed to any of the institutions with which they are affiliated.

1. The literature on counterfeiting and piracy is large and cannot be fully reviewed here. Instead, we focus on widely-cited studies of counterfeiting and piracy and offer an overview of the various approaches that researchers have adopted.
2. See, for example, Newton et al. (2010). Counterfeit drugs are part of the broader challenge of substandard pharmaceuticals, which also extends to generic medicines.
3. Taylor (1993) analyzed such “masking” strategies, which are costly in welfare terms. Again, Maskus (2000; 2012) offers anecdotal evidence of such responses in developing countries.
4. Copyright protection is limited to strike a balance between incentives for creating new works and the desirability of making them widely available. Thus, copyright has a limited term and certain
exceptions and limitations to exclusive rights exist in areas in which society reaps special benefits from the rapid dissemination of creative works.

5. Maskus (2000) relates anecdotal evidence that this price moderation occurred in the software industry when Taiwan improved its enforcement activities in the 1990s.

6. Conner (1995) offers a broader set of circumstances in which it might be profitable for copyright owners to accommodate copying in the presence of network effects and quality differences.


8. Another problem is that the analysis takes place at the highly aggregated two-digit Harmonized System categorization of traded goods. Within each such category, typically hundreds of goods subcategories may or may not be subject to much counterfeiting. The OECD (2008) analysis assumes that its average rates computed at the aggregate level apply to each subcategory. This list of methodological problems is not exhaustive. The OECD (2008) report describes several other important problems that likely bias the resulting estimates.

9. The difference between the total installed software and legitimately purchased software is the amount of piracy. Although this approach is sensible in that it relies on reported measures of legitimate usage in relation to overall computer capacity, both of which may be reliable, the BSA report does not indicate what the underlying sampling biases may be. This approach also ascribes some use of open-source software and freeware to piracy.

10. See, for example, Siwek (2007) and Frontier Economics (2009). The latter study considers the reemployment of workers “losing” their jobs as a result of counterfeiting and piracy but does not consider the effects of revenue-compensating tax policies. In addition, the assumption of a certain share of workers not finding reemployment implicitly assumes some form of historical shock that led to actual job losses. The latter assumption is difficult to reconcile with the more gradual evolution of counterfeiting and piracy activity observed in most sectors.

11. Maskus (2005) offers a simple partial-equilibrium model of the effects on employment and sales in several key sectors in Lebanon, using data for 1996. However, this approach cannot determine cross-industry equilibrium impacts and is only suggestive at best.


13. This outcome was noted in survey evidence of the pricing strategy of legitimate software producers in Asian countries as copyright enforcement improved (Maskus 2000).


15. See Fink (2013) for a discussion of how trade agreements treat intellectual property rights and how they may affect economic welfare.

16. There are also recent attempts to employ surveys to estimate the incidence and characteristics of counterfeit and substandard drugs. For example, Bate et al. (forthcoming) studied 1,437 samples of Ciprofloxacin in 18 low-to-middle-income countries. These authors found that 9.88% of samples contained less than 80% of the correct active ingredient, and 41.5% of these failures are counterfeits. Counterfeit and substandard drugs tend to differ in two observable attributes: first, counterfeit drugs are more likely to mimic drugs registered with local drug safety regulators than substandard drugs. Second, after controlling for other factors, substandard drugs are on average cheaper than generics, whereas counterfeit drugs do not show any significant price difference.

References


Measuring Violent Conflict in Micro-level Surveys: Current Practices and Methodological Challenges

Tilman Brück, Patricia Justino, Philip Verwimp, Alexandra Avdeenko, and Andrew Tedesco

This paper reviews current practices and common challenges in the measurement of the causes, functioning, and consequences of violent conflict at the micro-level. We review existing conflict- and violence-related survey questionnaires, with a particular focus on the World Bank’s Living Standard Measurement Surveys. We discuss methodological challenges associated with empirical work in conflict-affected and fragile areas—such as operationalizing a definition of conflict, using the appropriate units of analysis, deciding on the timing of the survey, dealing with data biases, and conducting surveys in an ethically sound manner under conditions of insecurity—and propose ways to improve the usefulness of existing surveys to analyze conflict processes at the micro-level. Violent conflict, households, survey methods, questionnaire design. JEL codes: C81, D74, F52, O12

Introduction

Analyzing the causes of violent conflict has preoccupied social scientists for a long time. The academic literature on conflict has traditionally had a strong macro perspective, with a focus on understanding the rise of violence against state institutions, and between different ethnic groups (Horowitz 1985; Skaperdas 1992; Collier and Hoeffler 2004; Fearon and Laitin 2003). This literature has led to considerable advances in our knowledge on the complex causes of political violence. It has, however, offered a more limited understanding of the role of the causal mechanisms and micro-level dynamics that may shape the relationship between violent conflict and social, economic, and political outcomes.
These concerns have resulted in a growing new research agenda on the micro-level analysis of violent conflict.\(^1\) This emerging body of research has begun to shed light on some of the complex micro-level causes and consequences of violent conflict by generating important theoretical and empirical insights on a number of dimensions of violent conflict processes (see Justino 2012 for a comprehensive review). But despite considerable progress, we still have limited rigorous and comparative evidence of how people live in contexts of violence and conflict, particularly regarding the following issues: what choices they make to secure lives and livelihoods; how institutional structures impact on and are affected by these decisions; or what policies work in establishing peace and supporting economic stability in areas and among populations affected by violent conflict.

This lack of systematic understanding is explained to a large extent by the limited amount of information we currently have on how people live and survive in areas of violent conflict. Over the last three decades The World Bank and other institutions have developed highly sophisticated surveying techniques, most notably the Living Standards Measurement Studies (LSMS), to collect socio-economic data at the micro-level. These advances in data collection have led to a wealth of knowledge on how individuals, households, and communities live and adapt to a variety of shocks and life events, including price changes, sudden climatic shocks and loss of work, or illness, among others. The impact of political shocks and events such as violent protests, communal riots, revolutions, civil wars, genocide, and international wars, is less-well understood. Several large household surveys conducted in conflict-affected countries such as Nigeria, Indonesia, Colombia, Rwanda, South Africa, Pakistan, and Liberia, only sporadically feature (a limited number of) questions that capture the effects of violence and other forms of political instability and fragility.

The absence of detailed information on how individuals, households, and communities experience violent conflict means that researchers typically rely on crude proxies of conflict (often whether dwellings have been destroyed, the number of deaths in the household and whether the households has ever been displaced), which makes it hard to build a systematic and comparable understanding of how processes of violence and insecurity have affected different people and communities, the channels through which violence may affect welfare and behavior at the micro-level, and how violent conflicts may transform societies, politics, and local economies.

The objectives of this paper are to review current empirical research on conflict processes at the micro-level, discuss the methods used to empirically capture such processes, and suggest potential advances to current survey methodologies. The paper is organized as follows. In the following section, we review current survey practices in conflict-affected contexts, paying close attention to recent academic literature that makes use of quasi-standardized institutional surveys, such as The World Bank’s LSMS. In the subsequent section we discuss recent studies based on surveys purposely designed to investigate conflict processes, and reflect on...
common methodological challenges related to the design, implementation, and analysis of micro survey data in conflict-affected contexts. The final section concludes by discussing ways forward to improve methodologies for data collection at the micro-level in conflict-affected contexts.

Current Empirical Approaches to Understanding Conflict at the Micro Level

Social scientists have used a variety of empirical methods to study the impact of violent conflict on human welfare and behavior. Recent advances have been possible thanks to the wider availability of rigorous evidence from conflict-affected contexts. This new empirical research has developed in two broad directions (see Verwimp, Justino, and Brück 2009, and Justino, Brück, and Verwimp 2013). The most common direction has been the use of socio-economic datasets in conflict-affected regions that were not explicitly collected for the analysis of processes or consequences of violent conflict per se, but either contain a number of variables (often self-reported) that can be used as proxies for human exposure to violent conflict, or can be creatively merged with conflict event data. The second direction is based on data specifically collected to identify the causes and functions of violent conflict at the micro-level. This is the ideal approach because it allows researchers to tailor the surveys to directly address important research questions about different aspects of conflict processes, their causes, and their consequences. It is, however, a less common approach due to the high costs of these surveys, the level of resources required, and the ethical and security constraints associated with doing primary research in areas of violence. In this section we review recent studies based on existing socio-economic surveys that were not conducted to analyze conflict processes, but have been used to generate important insights on conflict processes and its welfare outcomes. We focus on studies that have used the widely available LSMS conducted by the World Bank and partners, and national census data and the Demographic and Health Surveys (DHS). We also discuss these studies’ insights and propose ways in which the surveys could be improved in order to advance knowledge on conflict processes. We discuss the purposively designed survey approaches in the methodology section.

Using Living Standards Measurement Surveys to Understand Conflict Processes

The LSMS, implemented by the World Bank, aims to provide high quality data for policy makers to assess the effectiveness of interventions designed to improve the
living standards of individuals, households, and communities in countries where they are conducted. Micro-level empirical research on the effects of conflict and violence has made use of several of these surveys.

The main advantage of the LSMS is their comprehensive treatment of household welfare, and how they detail topics such as household demographics, income, health, labor, and education. However, these surveys are primarily designed to be conducted in peaceful contexts and often neglect to explicitly address violent conflict as a category in its own right, even when they are implemented in countries affected by conflict. The LSMS are also designed to meet the needs of governments, who sometimes wish to avoid referring to the conflict in an effort to start afresh. As a result, questionnaires may focus on the experiences and the standards of living after, rather than during, the conflict. Some questions about conflict may also be politically sensitive and government officials may be apprehensive about including them in household surveys. For example, questions about the destruction or theft of assets that identify the perpetrators, especially if government forces are included in the list, may raise controversial or even legal issues for government administrations. Likewise, government-sponsored surveys may avoid addressing the conflict in formerly rebel-held territories for fear of invoking distrust or upsetting a delicate peace settlement. As a result, some questions that are relevant for researchers may be left out in government-sponsored questionnaires.

Several LSMS have nonetheless included conflict-related questions and yielded important insights for conflict research. In other cases, researchers have been able to match the information in these surveys to external conflict event datasets in order to design identification strategies that allow the causal analysis of conflict effects on individuals, households, and communities. For this paper we reviewed 24 of these surveys and analyzed their structure and content. In general, we found that including conflict questions in LSMS has been done in a piecemeal fashion, resulting in insights scattered across countries and categories rather than a systematic and comparative approach to measuring conflict. But we also found good examples of conflict-sensitive questions across these surveys that have resulted in important new knowledge on the impact of conflict on the lives and livelihoods of people in areas of violence. These studies have led to significant advances in our understanding of the causal impact of violent conflict on key aspects of people’s lives, including dynamics of poverty and food security, education and health, livelihood choices, labor supply decisions, intra-household decisions, and individual, household, and group behavior. We provide a summary of existing socio-economic surveys with conflict related questions in Appendix I, and in Appendix II summarize studies (and their main findings) that have used LSMS (and DHS) to understand conflict processes.
Using other Standardized Household Surveys and National Census Data

In addition to the LSMS, researchers have leveraged other standardized household surveys and national census data to study conflict processes. For example, Deininger (2003) conducted one of the first micro-level analyses of violent conflict and its consequences using data on communities and households included the 1999–2000 Uganda National Household Survey (UNHS) and the 1992 Uganda Integrated Household Survey (IHS). These surveys contain information on approximately 10,000 households and 1,000 communities. The surveys asked respondents questions related to the civil war, including information on victimization and motivations for participation in the war. Czaika and Kis-Katos (2009) have studied the determinants of displacement in Aceh, Indonesia, using the Village Potential Census (PODES), which maps conflict-affected villages across the whole of Indonesia. The census itself includes questions, posed to community leaders, on conflict events in particular communities. Rohner, Thoenig and Zilibotti (2012) have matched two waves of the Afrobarometer survey, a widely available socio-economic survey, to the ACLED dataset in order to investigate the effect of the conflict in Northern Uganda on social capital.4,5

Verpoorten (2011) demonstrates how to use widely available census data to indirectly measure conflict mortality. Census data is generally quite comprehensive and includes mortality information on victims targeted by all combatant parties, those dying in both large and small events, those in both remote and accessible areas, as well as direct and indirect mortality levels. Census data has also been used in Weidmann (2009), who combined the Bosnian census with the ACLED conflict database to determine how conflict affects ethnic population concentration.

Some researchers have uncovered important insights into the human capital effects of violent conflict through the creative use of historical datasets. One good example is Akbulut-Yuksel (2009), which shows how a unique data-set on city-level destruction in Germany caused by Allied Air Forces bombing during World War II can provide far-reaching insights when combined with a socio-economic panel. This is one of the first studies to capture the long-term human capital effects of violent conflict across several generations. Historical archives and surveys of survivors have been used recently by political scientists to examine the long-term social and political legacies of the internal conflict in Greece in the 1940s (Kalyvas 2006) and the Spanish Civil War (Balcells 2012).

Using Demographic and Health Surveys (DHS) to Identify the Human Capital Effects of Violent Conflict

Demographic and Health Surveys (DHS) are specialized surveys designed to monitor health, fertility, and mortality outcomes in several developing countries. Although

Brück et al.
these surveys often lack information on conflict and violence—even when conducted in conflict-affected countries—several studies have leveraged them to analyze the demographic, health, and education effects of violent conflict. The methodologies followed are similar to those used with the LSMS. For instance, Bundervoet (2009) investigated the profile of victims of the 1993 killings in Burundi using the United Nations Population Fund (UNFPA) demographic survey conducted in Burundi in 2002. Verwimp and Van Bavel (2014) have used the same survey to study the effect of conflict in Burundi on the gender-gap in primary school completion. De Walque and Verwimp (2010) used Rwandan DHS data from 2000 and 1992 to estimate excess mortality in the 1994 genocide. De Walque (2004) made use of the DHS to assess the long-term impacts of the Cambodian genocide during the Khmer Rouge period. Schindler and Brück (2011) used DHS data to assess the effects of the genocide on (replacement) fertility in Rwanda. The potential of DHS in conflict research remains, however, largely underexploited.

Responding to Challenges in the Use and Design of Standard Surveys

Researchers have made use of existing socio-economic datasets collected for purposes other than conflict research by creatively using occasional questions related to conflict or by merging them with conflict event data. This approach makes good use of existing data and has provided valuable insights on the effect of conflict on socio-economic outcomes at the micro level. However, these studies face important shortcomings. Most notably, studies that rely on existing surveys that were not purposively designed to understand conflict processes—as is the case with most of the current literature—are generally based on an ad-hoc treatment of conflict, typically based on a limited number of self-reported indicators of household exposure to violence.

These shortcomings can be mitigated through the use of high-quality event datasets, as demonstrated in several studies discussed above. The main downside of this approach is that matching datasets may be difficult, either because names of locations or identities of respondents cannot be reconstructed, or because researchers are not able to access that information for confidentiality reasons. However, given the many security, financial, and human resources trade-offs faced by researchers working on violent conflict at the micro-level, large institutionalized socio-economic surveys will often be the best option for gathering conflict research information. A question remains: Can these surveys be improved in ways that will advance research and knowledge on conflict processes at the micro-level?

Below we suggest a number of guidelines to improve the sensitivity of future standard socio-economic surveys conducted in conflict-affected contexts. We propose four ways in which these surveys could be adapted to capture more accurately
conflict-related events: (i) allow respondents to self-report on conflict events more comprehensively by including conflict-related scenarios in answer categories; (ii) record the timing of events; (iii) be sensitive to the type and intensity of violence; and (iv) include conflict questions across several survey sections and include a range of conflict-related choices in answer categories. We have used these guidelines elsewhere to construct the “Conflict Exposure Module” to be inserted in existing socio-economic surveys conducted in conflict-affected countries.6

**Respondent Self-reporting**

Asking respondents to self-report on how conflict has affected them is a straightforward way to learn more about conflict dynamics. Too few existing socio-economic surveys extend answer choices to give respondents the opportunity to explain how different dimensions of the conflict may have affected them. For example, the LSMS conducted in Tajikistan 2007 asked respondents why they did not work in the past 14 and 30 days. This question could be modified to provide a better understanding of conflict effects by including answer options such as “lack of security” or “disability due to violence”. The 2006 Iraq LSMS is one example where the effect of conflict on income is addressed by including answer options that point to “security” and “handicaps” as reasons for inability to work. These categories could be made more comprehensive by including a broader range of conflict-related scenarios, such as discrimination, crime, destruction of assets, disappearance of key markets, or military service.

The use of self-reported information has experienced setbacks. Self-reported answers are highly subjective and may introduce biases because respondents may erroneously recall the reasons why they made certain decisions. In situations as intense as conflict, respondents may construct a narrative that gives disproportionate weight to extraordinary and memorable experiences. Respondents may also be unable to distinguish between ultimate and proximate causes. Despite these potential limitations, including conflict-related scenarios in the answer categories of questions routinely asked in large socio-economic surveys would be a cost-efficient way of adapting existing surveys to better understanding conflict contexts. Some of these issues may be solved through econometric techniques such as the use of instrumental variables.

**Sensitivity to the Timing of Events**

In several of the studies we reviewed for this paper, carefully recorded information on when events occurred has allowed researchers to match socio-economic outcomes to conflict events. However, few socio-economic surveys in conflict-affected areas record time information systematically. Detailed time information would allow researchers to know whether events occurred before, after, or during a
conflict, and therefore capitalize on conflict event databases that provide a localized history of conflict events. Examples include time information on when household members left or joined their families, when income, asset, and food consumption losses occurred, when coping measures were introduced, when harm was inflicted, and when people were displaced or migrated.

Researchers may also be able to construct localized conflict timelines by recording dates of major conflict events and allowing respondents to describe when an event occurred in reference to these timelines. This usually requires researchers to establish a local conflict event timeline beforehand, preferably with the aid of an initial community questionnaire. Respondents would then be able to remember whether specific events in their lives may have occurred in relation to, for instance, major attacks or other easily recalled events. The overarching goal of being time sensitive is to align respondents’ answers to major changes in the conflict’s dynamics, rather than produce a precise chronology. Existing socio-economic surveys can be adapted without adding substantial extra costs by introducing questions that capture the timing of events.

**Sensitivity to the Type and Intensity of Violence**

Micro-level surveys in conflict-affected areas could be more sensitive to different modalities of violence that take place during violent conflicts. Each conflict creates its own hardships, which local populations feel acutely. Measuring how much people suffer is important for determining the sources of acute hardship and the type of violence borne by the population. For example, respondents in Angola suffered physical injuries from the widespread use of land mines, while Palestinians have suffered losses of income from the difficulty of moving through checkpoints. Further, Congolese families are ripped apart because of the widespread use of rape and sexual violence. Sensitivity to the type and intensity of violence requires answer categories to go beyond simple binary variables by including a range of responses at different levels of intensity.

**Comprehensiveness**

Overall, surveys could be more comprehensive in relation to conflict-related events by covering a broad range of channels whereby conflict may affect both lives and livelihoods. While LSMS are by their nature very comprehensive, conflict-sensitive questions are often left out of many modules. Surveys that focus too narrowly on select categories may fail to account for the multi-dimensional impact of violent conflicts. In order to address this issue, socio-economic surveys could include conflict-related questions across their various sections on demographics, economic welfare, coping
activities, health, migration, education, perceptions of security, life satisfaction, and
expectations. Not all existing socio-economic surveys will be able to follow these guidelines at
all times given the number of financial, political, and logistical tradeoffs they may
entail. However, more investment in the development of appropriate methodological
systems may lead to considerable advances in gathering rigorous, systematic, and
comparable evidence across different conflict-affected contexts. Better knowledge
will, in turn, result in better and more effective policy interventions to provide
physical and economic security to the millions of men, women, and children that
continue to live in persistent cycles of violence and conflict.

Recent Methodological Advances for Surveying
in Conflict-affected Areas

Purposely designed studies make up an emerging body of research on the causes
and impacts of conflict at the micro level. These studies’ great advantage is that
they identify and measure conflict directly within the survey questionnaire, allowing
researchers to identify more precisely the complex causal mechanisms that
shape the relationship between violent conflict and individual, household, and
community welfare and behavior. We review here examples from five types of purposely
built surveys used in the literature: (i) ex-combatant surveys; (ii) genocide and
atrocities surveys; (iii) surveys of displaced populations; (iv) post-conflict reconstruc-
tion surveys; and (v) conflict surveys conducted among civilian populations.

Ex-combatant Surveys

Several surveys have been conducted to analyze the experiences of specific popula-
tion groups living through violent conflict, notably former soldiers and members of
rebel movements. Below we discuss four surveys that have developed novel and creative instruments to capture processes of recruitment, armed group support, and
combatant-civilian relations.

The Surveys of War Affected Youth (SWAY) in Northern Uganda, conducted by
Chris Blattman and Jeannie Annan in 2005 and 2006, have made an important
contribution to research on the micro-level effects of violent conflict. The surveys
implemented a novel sampling methodology—employing a technique which the
authors define as “retrospective sampling”—that attempted to reconstruct the
sample before the conflict took place. This helped the researchers correct for attri-
tion bias resulting from conflict-related deaths or migration. The study also
creatively used culturally specific indicators to measure psychosocial well-being, such as “nightmares and insomnia” or “perceptions of haunting by spirits”. Although culturally-specific, introducing these questions has provided important directions for the collection of hard-to-quantify indicators in survey instruments.

Humphreys and Weinstein (2004; 2008) conducted a novel survey on the motivations of Sierra Leonean ex-combatants for joining and staying with armed groups, and their attitudes towards disarmament, demobilization, and reintegration. In particular, the authors were able to reconstruct time periods within the conflict by asking respondents to recall their geographic location during active participation in armed groups at specific periods during the war. This allowed them to analyze both processes and motivations of recruitment across time. In order to assist recall, the authors constructed a timeline of well-known events and dates and placed individual answers within these different time periods. Another innovative aspect of this survey is the way in which sensitive questions were phrased. In order to avoid respondents feeling compromised by their answers, the survey focused on asking whether respondents observed potentially incriminating events (such as theft, rape, and assault), rather than asking whether respondents perpetrated them personally. Overall, contrary to established wisdom that portrays ex-combatants in Sierra Leone as motivated by greed and the gains of looting, the survey finds evidence for an array of motivations for joining and staying in armed groups, ranging from fear and forced recruitment, to access to protection and the provision of basic needs.

Arjona and Kalyvas (2008) also studied individual motivations for joining armed groups in Colombia; their survey was conducted among 732 ex-combatants of a leftist guerrilla group and a right-wing paramilitary group. The sampling approach followed in this survey is more challenging than the survey conducted in Sierra Leone due to security concerns in various sampling areas, and the reliance on sample frames drawn exclusively from a national demobilization and reintegration program. These factors have prevented the authors from constructing a representative sample. However, the survey is unique in that it offers extensive information on individual motivations for joining armed groups, as well as on how different armed groups are organized and their relationship with civilian populations, which has allowed a better understanding of how armed groups exercise control of different populations, and how local populations adjust their behavior in the presence of armed groups (see Arjona 2010).

Mvukiyehe, Samii, and Taylor conducted over 3,000 interviews in 2007 among ex-combatants in Burundi, primarily focusing on armed group recruitment. This survey was implemented among combatants and non-combatants in order to identify how experiences of violence differed between the two groups. The survey covers personal experiences of violence, which have not been captured elsewhere, such as physical mistreatment, sexual abuse, or forced labor. This information is important for the design of policy interventions that consider the specific war experiences of combatants.
Genocide and Atrocities Surveys

Genocides are extreme violent events that produce enormous welfare impacts and pose unique challenges for researchers. In this section, we discuss two studies: the Genocide Transition Survey (2000) and the Darfur Refugee Questionnaire (2006).

The Genocide Transition Survey (2000), conducted in Rwanda by Verwimp, is one of the first examples of the potential for panel survey-based research in conflict-affected areas. Verwimp tracked the fate of household members who had been interviewed in a nationwide agricultural survey prior to the 1994 genocide. In addition to important insights into the profiles of perpetrators (Verwimp 2005) and victims (Verwimp 2003b) of the 1994 genocide in Rwanda, this survey has demonstrated that tracking households and individuals is possible even under the difficult circumstances of such extreme violent events. By creating a systematic profile of genocide perpetrators, the survey presented a more nuanced portrayal of the genocide in Rwanda beyond the simple discourse about ethnic rivalry between Hutu and Tutsi. The study particularly showed how economic considerations and issues around land distribution may have affected the outbreak of the violence as much as pre-existing ethnic rivalries.

The Darfur Refugee Questionnaire (DRQ) laid the foundations for the U.S. State Department to declare the killings in Darfur as genocide. The survey solicited a description of violent acts from the victims surviving in refugee camps and linked them to their perpetrators. Details on this complex survey are provided in Totten and Markusen (2006). This survey has been critical in policy interventions aimed at the return and reintegration of displaced populations, and the design of political settlements that may avoid the re-ignition of the conflict.

Surveys of Displaced Populations

The welfare losses suffered by displaced persons are an important area of research in conflict studies. Two of the most prominent surveys of displaced populations are the Northern Uganda Livelihood Survey (NULS; 2007) and the Deininger, Ibáñez, and Querubín (2004) study in Colombia.

The NULS was conducted in 2007 (see Bjørkhaug et al. 2008). The survey is a follow-up of the 2005 Northern Uganda Internally Displaced Persons Profiling Study and the 2006 Lira District Early Recovery Needs Assessments conducted by Fafo in Norway. The carefully phrased questions in NULS are specific enough to capture important changes in people’s lives due to violent conflict, including motivations for migration, experience of violent crime and abduction, other forms of victimization and causes of health problems, as well as expectations of the future. Additionally, the survey identifies whether the person was a combatant, thus providing important
information on how combatants and displaced civilians may experience violence. Studies using the NULS 2007 include Bozzoli, Brück, and Muhumuza (2011 and 2012).

Deininger, Ibáñez, and Querubin (2004) use an unusually large survey conducted among 32,093 households applying for assistance from the Catholic Church in Colombia to investigate the decision to return after displacement. The survey contains unique information on the causes of displacement, household demographics, access to land, and labor market and education outcomes. This is one of the few surveys available that trace the movements of displaced people. Information was collected only if people requested assistance from the church, which may lead to some selection biases. However, this information has been used insightfully to examine the extent of asset losses and labor market prospects of displaced people (Ibáñez and Moya 2009), the determinants of displacement (Engel and Ibáñez 2007), and labor supply outcomes and wage changes for displaced people (Calderón and Ibáñez 2009).

These two surveys have also been paramount in questioning some of the assumptions underlying return policies for refugee and displaced populations. Notably, the two surveys have highlighted the fact that many population groups may not desire to return to original locations for a variety of reasons ranging from fear of violence, to the fact that they have, over the course of displacement, built new lives in their new locations that they do not want to abandon (Bozzoli et al. 2013). These findings are at odds with international efforts to relocate displaced and refugee people, and raise important implications in terms of reconciling the needs of displaced and host communities.

Post-conflict Reconstruction Surveys

Some surveys have been developed by international institutions operating in conflict zones to assess the sustainability of post-conflict reconstruction. One example is the Standardized Monitoring and Assessment of Relief and Transitions (SMART) surveys, which provide a standardized methodology for measuring key statistics in the wake of complex emergencies (including conflict). The method focuses on basic indicators such as the nutritional status of children under 5 and mortality rates.\(^\text{12}\)

The International Committee of the Red Cross (ICRC) and the Greenberg Research team conduct the “People on War Surveys” in a variety of conflict-affected countries (ICRC 2009).\(^\text{13}\) The ICRC funds the surveys in part to assess the perception of its own interventions. The surveys are standardized so that results can be compared across all participating countries. To account for country-specific contexts, the wording of some questions is modified where necessary.

Mvukiyehe and Samii (2008 and 2009) have conducted a survey to evaluate peacekeeping operations in the Cote d’Ivoire. This survey captures the potential for
Conflict re-escalation by reporting on events and circumstances that might warn of renewed conflict. The survey also investigates perceptions of security amongst populations and repeated violence against civilians in different locations. Similar to the Humphreys and Weinstein ex-combatant survey discussed above, respondents were asked whether or not they witnessed or suspected “inter-ethnic fighting, presence of armed groups, or recruitment by armed groups in their localities” in relation to time periods constructed between well-known events, which has allowed researchers to assess how early conflict signs have evolved across time.

Conflict Surveys Conducted among Civilian Populations

Several socio-economic household surveys have incorporated comprehensive components that capture the effect of various forms of violent conflict on civilian respondents. The Burundi Priority Household Panel (1998, 2007, and 2012) was designed to provide detailed information on the welfare effects of the civil war in Burundi by comparing households in villages affected by the war with households in non-affected areas (see Bundervoet et al. 2009). The survey features questions on violence and conflict at the individual, household, and community levels. The panel design, collected in three waves in 1998, 2007, and 2012, captures comparable data on welfare before and after incidences of violence. Special attention was given to tracking individuals who left the household since the first wave of the survey (Verwimp and Bundervoet 2009). The same team followed up the results of this survey with experimental behavioral games in conflict-affected and non-affected areas in 2009 (Voors et al. 2012). This unique setup has allowed the researchers to link outcomes measured in the survey with behavioral data, thereby proving important insights into how exposure to violent conflict may affect people’s fundamental preferences and pro-social behavior.

The Life in Kyrgyzstan Survey (LIK) interviewed 3,000 households annually over 3 years (2010 to 2012) to create a nationally representative panel (Brück et al. 2014). The LIK surveys cover a comprehensive list of topics, following similar protocols to the LSMS, with the additional inclusion of modules designed to elicit information about local tensions and conflicts, which has allowed researchers to better understand how local conflicts may affect development outcomes.

The Maharashtra Household Longitudinal Survey (MHLS) was conducted in 2010 by the Institute of Development Studies (UK) among 1,089 households living in violence-affected areas in the Indian state of Maharashtra (see Gupte, Justino, and Tranchant 2012). The same households were surveyed again in 2012. This unique panel study was designed to capture how individuals and households live in areas characterized by persistence communal violence in India. This survey has allowed
researchers to understand for the first time the intersection between economic vulnerabilities and exposure to persistent forms of urban violence in slum areas in India.

The Colombian Longitudinal Survey of Wealth, Income, Labor, and Land (ELCA) interviewed 10,000 households in rural and urban areas affected by conflict in Colombia. The first wave was conducted in 2010 and the second wave took place in 2013. This is the most comprehensive longitudinal survey ever conducted among conflict-affected populations. A novel aspect of the survey is the questionnaire modules on activities of armed groups in different neighborhoods (see Gafaro, Ibanez, and Justino, forthcoming), as well as detailed information on migration, recruitment, and local cooperation with armed groups. These surveys allow for the first time a comprehensive understanding of how violent conflict in its various modalities may lead to specific changes in institutional arrangements and individual and group behavior that may explain the persistence of violence in some areas of Colombia, but not others, with profound implications for the ongoing peace process in Colombia, as well as conflict resolution interventions elsewhere in the world.

Responding to Challenges in Purposely Designed Surveys

Purposely designed surveys conducted in conflict-affected areas are the state-of-the-art of empirical conflict research at the micro-level. There are, however, only a limited number of these surveys due to a number of challenges. First, primary fieldwork research in conflict-affected countries is quite expensive due to a lack of infrastructure and difficulties in engaging with suitably qualified local field researchers. Second, security concerns can also pose difficulties to the research team and their respondents. In addition, researchers face a series of methodological challenges. Some of the most prevalent challenges include the following: (i) defining conflict at the micro-level; (ii) choosing the appropriate unit of analysis; (iii) identifying time dimensions in survey questionnaires; (iv) dealing with data biases (such as selection bias and recall error); and (v) addressing ethical and security issues associated with doing primary research in conflict-affected contexts. We discuss these common challenges below.

Defining Conflict at the Micro-level

One of the most important challenges to designing surveys in conflict-affected contexts is to create and operationalize a definition of conflict that captures the complex impact of conflict on the lives of individuals, households, and communities. Several authors have proposed more or less overlapping typologies of violent conflict (Gupta 1990; Singer and Small 1994; Sambanis 2001). These definitions are useful for understanding conflict as a macro-phenomenon, but are difficult to
uphold at the micro level because they are too far removed from the everyday disturbances experienced by local populations (Justino, Brück, and Verwimp 2013).

An additional difficulty in defining violent conflict from a micro-level perspective is to determine when a violent conflict starts and ends. In particular, lulls or spikes in violence may make the conflict feel as if it starts and stops rather than persists at a continuous intensity. Even after a conflict has subsided at the national level, the persistence of lower levels of violence and instability may continue to affect households and their members. Likewise, as conflicts draw to a close, changes in the identity of the belligerents may create new coping dilemmas for the population. Many individuals and groups living in conflict-affected areas find themselves, therefore, responding, acting, and being affected by stages in between conflict and peace. Macro-level concepts of time periods may miss these nuanced variations at the micro-level. As conflicts change frequently over place, time, and context, it is necessary to have a broad definition of conflict while also establishing observable characteristics that can be easily captured through empirical data collection.

**The Unit of Analysis**

The second methodological point in the design of surveys in conflict-affected contexts is the choice of the appropriate unit of analysis for different types of questions. Depending on the type of information sought, survey questions typically target individuals, household heads, or central figures in the community. The individual is the lowest level of analysis. The objective of using individual-level questions is to determine how individual decision-making—across gender, age and different socio-economic backgrounds—may respond to the impact of violence on livelihoods, well-being, and security. Individual-level surveys may also be able to capture specific individuals that may have been directly involved in the conflict, such as soldiers, refugees, and displaced people.

In household-level surveys, the head or another member of household responds on behalf of the household. Household-level questions allow the assessment of the impact of shocks on households and their reactions as collective decision makers. Questions can target changes in access to services, markets, investments, and land, which may affect the entire household even if only a few members are directly involved. Household-level questions can, in addition, be used to draw a broader picture about social relations and networks.

Community-level questions may also be useful for uncovering the extent of the impact of violent conflict, especially when violent conflict events are concentrated in time and space. Community-level survey components may be able to generate a conflict history that records the overall characteristics of localized events. This information can be used as a starting point for designing household and individual-level surveys and, crucially, to provide context when gathering time information. Community-level
analysis may also allow for a more accurate determination of deaths across the community, for instance by examining listings of names in local administrative records. Moreover, knowledgeable members of communities often provide important qualitative and quantitative insights. Community-level questions are also useful for assessing migration flows, urgent needs of the community, and the impact of policy interventions, particularly reconstruction interventions in the post-conflict period.

**Time Dimension**

The timing of surveys is decisive for the quality of subsequent analysis. As with most socio-economic data, the quality of people’s responses about conflict processes tends to diminish as the time between the conflict and survey widens. Yet in many instances, the intensity of conflict experiences may facilitate memory, and data collection in conflict-affected areas may be done well after a conflict has ended. However, much depends on the circumstances of the conflict and its long-term legacies. Researchers have used three main approaches to survey timing, including administering the survey while the conflict is ongoing, administering the survey *ex post* by asking respondents to assess before and after conditions, and making use of panel data.

If the violent conflict is still ongoing at the time of the survey, researchers tend to use a 12-month reference period to elicit information on the short-term effects of violent conflict on individuals and households. This reference period has a number of advantages. First, its frequent use in other socio-economic surveys may allow for comparability. Second, this time period is useful for gathering economic data that may contain seasonal effects, such as any indicator linked to agricultural or climatic cycles. However, before employing a 12-month reference period, researchers should ensure that this is appropriate given the dynamics of the conflict. It may be that in the last 12 months, or in the period immediately before, a major conflict event significantly affected the respondents. In these cases, it may be better to refer to the conflict event specifically in the reference period. Researchers may employ variants of prompting devices such as “since your village was attacked,” “since the beginning of the conflict,” or “since armed fighting ceased in your area.”

Surveys cannot always be conducted very close to conflict events. When conducting a survey some time after a conflict, researchers have addressed the issue of temporal comparison by asking respondents to recall aspects of their lives before and after the conflict. Many questions asked in existing surveys address the problems of missing *ex ante* data by regularly using phrases such as “before the conflict” or “since the start of the conflict”. These types of questions can create further time variation by asking respondents to recall living standards at specific points during the conflict, usually demarcated by well-known events. Humphrey and Weinstein’s (2008) work in Sierra Leone provides a good example of this technique. Such *ex post* surveys may, however, introduce potentially severe biases, as respondents...
may erroneously recall events, overestimate their levels of welfare before the conflict, or samples may exclude important sub-groups. These surveys must therefore pay particular attention to the design of reliable timelines of events and other mechanisms to elicit accurate information from respondents’ memories.

Researchers may be able to collect longitudinal data if they are fortunate enough to have access to a survey that was done before, and reasonably close to, the onset of conflict. Panel datasets offer rich time variation and minimize many of the concerns about biases prevalent in other methods. However, follow-up surveys must be especially careful to control for attrition. People in conflict-affected areas tend to be highly mobile and subject to a high degree of mortality, making them difficult or impossible to include in the follow-up survey. When these groups systematically differ from the overall population, excluding them will bias the sample. However, recent work discussed above in Burundi, Rwanda, Indonesia, India, and Colombia, among others, has shown that reliable longitudinal data can be effectively collected in conflict-affected countries.

Data-related Biases

Research in conflict-affected areas takes place under unusual and often insecure circumstances, adding extra difficulties for researchers attempting to create a representative sample. This may even continue into the post-conflict period as governments may exclude researchers from sensitive areas, such as recently in Sri Lanka, Egypt, Syria, or Mali. Entire areas may be inaccessible, forcing researchers to rely on ex post surveys. Respondents may also self-censor answers to avoid any risk of retribution from insecure political authorities. Conflict researchers have employed several strategies to minimize such biases. Below we discuss some of the most prevalent biases in conflict contexts: selection biases and recall or response errors.

Selection Biases

Selection biases occur when individuals, households, or groups of households with particular characteristics cannot be sampled or interviewed. For example, declining economic activity during a conflict may result in entrepreneurial individuals migrating out, changing the characteristics of the population left behind. Or combatants may target specific ethnic and social groups during the conflict, forcing targeted populations to migrate or leading to a large number of deaths among specific population groups. In addition, panel datasets may suffer from attrition bias, as discussed above. The SWAY surveys addressed this issue by asking close family members to respond to an “absentee survey” on behalf of the respondent when tracking was impossible. This and other methods can be employed to help researchers generate a sample that has an equal likelihood of including individuals or households who have died or...
migrated out of a conflict-affected area. If respondents have died, tracking becomes impossible. Even in these circumstances researchers must take account of the potential biases created by attrition (see Justino and Verwimp 2012).

Recall and Response Errors

The length of recall periods has been a topic of discussion in socio-economic, demographic, and epidemiological surveys for a long time (Deaton 2001). Exposure to conflict may sometimes aid recall due the intensity of the event. However, the reverse scenario is also possible when respondents may repress traumatic memories or even refuse to talk about them. Training and sensitization of survey enumerators can alert them to the potential for these biases. As discussed above, recall can be aided by using event timelines that stimulate the respondent’s memory and accurately situate personal events in time. Other forms of response errors may take place in surveys conducted in conflict-affected contexts. For instance, if the survey sponsor is viewed with distrust or suspicion, respondents may provide misleading answers to questions or even refuse to cooperate entirely. In some post-conflict situations, government-sponsored surveys may run this risk, especially in formerly rebel-held territories. This may create biases if questions about conflict are included, as responses from those most affected may be inaccurate or absent. In the extreme, lingering animosity towards the government might be so strong that discussing conflict could unsettle the fragile peace and put the security of survey workers into jeopardy. For example, in the LSMS survey in Guatemala in 2000, the authorities asked relatively few conflict questions, though the conflict had only recently ended, because they feared that the population in formal rebel territories would refuse to participate. In post-genocide Rwanda, the government does not allow researchers to ask questions about ethnicity in surveys, thus limiting the capacity of researchers to link ethnicity with other variables of interest such as poverty, displacement, or gender.

Ethical and Security Considerations

Conflict surveys risk asking questions that may do harm to respondents. Researchers have a duty to weigh important ethical considerations while designing and implementing surveys. Sensitive questions may evoke traumatic memories about suffering, remorse, victimization, or guilt, potentially “re-traumatizing” respondents and harming them psychologically. Some questions may also risk incriminating or inviting retribution upon a respondent. If answers inadvertently become public, responses that identify perpetrators, victims, or actions taken by former combatants are particularly susceptible to this risk. Several mechanisms have proven useful to address and minimize potential ethical risks. The first and simplest way is to avoid asking some of
these questions. Researchers should be self-critical about whether their questions are strictly necessary, potentially harmful, or if there are less risky ways of obtaining the same information. To limit the risk of harm, it is also generally good practice to ask about group behavior rather than asking for specific names of perpetrators.

The training of enumerators and a local research team is a crucial aspect of doing high-quality surveys in sensitive and insecure areas. Ethics training should also provide comprehensive information about the risks that respondents face, and on adequate security responses. Proper training should also make enumerators aware of the emotional stresses associated with addressing sensitive issues surrounding conflict. A well-designed consent script is also important. Researchers should also be ready to respond should a harmful event occur, or should respondents become unduly distressed.

Final Remarks and Further Advances in Conflict Data Collection

We discussed in this paper how empirical research on conflict processes at the micro-level has flourished over the last few years due to the wider availability of good quality surveys conducted in conflict-affected countries. Some of the most comprehensive insights into the causes and consequences of violent conflict at the micro-level have been generated through purposely designed surveys. These surveys have been used to uncover unfolding process of conflict, rather than assessing conflict as a one-off shock, because they are able to collect systematic information on the various channels whereby different forms of violence may affect individuals and households. Well-designed surveys also provide valuable disaggregated information on conflict processes across time and place. However, these surveys require a lot of resources: sample sizes are large, interviews sometimes last several hours, transportation costs are high, and local expertise is crucial. Insecurity may also still be high in many of the areas being surveyed.

One way of minimizing these costs is to rely on existing large socio-economic surveys conducted by The World Bank and other international institutions, which may either contain self-reported information on conflict exposure, or can be matched to conflict-event datasets. This approach is less ideal due to the lack of precise information on the complex facets of conflict in each case study, as well as difficulties in comparing information across countries. Empirical work based on existing socio-economic surveys has nonetheless resulted in considerable advances in our current knowledge on violent conflict at the micro-level due to the development of a wealth of creative research methods to analyze conflict processes among violence-affected populations and across time. Although information on conflict is sometimes limited, the use of existing socio-economic surveys has many advantages over the use of purposely designed surveys, notably its costs and ease of availability. We suggested potential solutions for how these surveys can be made more

Brück et al.
comprehensive in a conflict-sensitive manner. We hope that these suggestions and
the wide discussion on various survey methods offered by this paper will lead to
further and better efforts to collect much needed information on the lives of individ-
uals, households, and communities affected by violent conflicts across the world.

Notes

Tilman Brück is with IGZ - Leibniz Institute of Vegetable and Ornamental Crops, Großbeeren,
Germany. ISDC - International Security and Development Center, Berlin, Germany, and the
Households in Conflict Network. Correspondence may be sent to: tilman.brueck@isd-center.org.
Patricia Justino is with the Institute of Development Studies, UK. Philip Verwimp is with Ecares, Centre
Emile Bernheim and Solvay Brussels School of Economics and Management, Université Libre de
Bruxelles, Belgium. Alexandra Avdeenko is with Mannheim University. Andrew Tedesco is an indepen-
dent consultant. The authors would like to thank Raka Banerjee, Kristen Himelein, Gary Milante, and
Diane Steele for their helpful comments and support. This paper is a shorter version of a sourcebook
on conflict measurement using household surveys developed for the World Bank (Brück et al. 2013).

2. For example, the LSMS conducted in Timor-Leste in 2001 (after the country had experienced
intense violent events in 1999) asks only two questions on war damage, focusing almost exclusively
on damage to dwellings. The LSMS conducted in Tajikistan in 2003 shortly after the Tajik civil war
asks only one question about war damage.
3. These include Azerbaijan (1995), four waves in Bosnia & Herzegovina (2001–2004), Guatemala
and Malawi (2004, 2010). For an excellent description of the design and use of the LSMS, refer to
Deaton (2000).
4. The Armed Conflict Location and Event Dataset (ACLED), is a comprehensive dataset on political
violence, including information on specific dates and locations of political violence, types of event,
groups involved, fatalities, and changes in territorial control.
5. De Luca and Verpoorten (2011) have conducted a related study using the same datasets.
7. Concrete examples are provided in: http://www.ids.ac.uk/files/dmfile/Measuring_Conflict_in_
Micro_Level_Surveys.pdf.
9. See also Fearon, Humphreys and Weinstein (2009), and the Sierra Leone PRIDE/JCTJ (2002). A
useful website for the collection of information on different surveys is the Post-Conflict and Ex-
10. A similar methodology has been adopted by Guichaoua (2007) to examine motivations to join
insurgent and incumbent groups in Nigeria.
11. See Verwimp (2003a) for a description of the survey.
12. The Centre for Research on the Epidemiology of Disasters compiles the Complex Emergencies
Database (CE-DAT), which includes SMART survey data.
13. Afghanistan, Colombia, Democratic Republic of the Congo, Georgia, Haiti, Lebanon, Liberia,
and the Philippines.
14. For a discussion of methodology in practice, see Green and Tony (2008), Verwimp and
Bundervoet (2009), and Bundervoet et al. (2009).
15. Two American IRB-approved ethics training exercises can be found at https://www.
citiprogram.org and http://phrp.nihtraining.com/. MICROCON, a large integrated program on conflict
research funded under the EU’s 6th Framework, developed guidelines for the ethical review of all data
collection efforts prior to field work. See www.microconflict.eu.
### Appendix I: Surveys with Conflict-related Questions by Regions

<table>
<thead>
<tr>
<th>Region</th>
<th>Survey Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angola</td>
<td>“Demographic and Health Survey (DHS).” 2006/7, 2011.</td>
</tr>
<tr>
<td>Chad</td>
<td>Demographic and Health Survey. 2004.</td>
</tr>
<tr>
<td>Eritrea</td>
<td>Demographic and Health Survey (DHS). 2002.</td>
</tr>
</tbody>
</table>
Appendix I: Continued

Iraq

Kosovo

Liberia


Demographic and Health Survey (DHS), 2007–2011 (3 surveys).

Malawi

Demographic and Health Survey (DHS), 1996–2010 (4 surveys).

Mali

Mozambique

Nepal

Nigeria
**Appendix I: Continued**

<table>
<thead>
<tr>
<th>Country</th>
<th>Survey/Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Demographic and Health Survey (DHS). 2008.</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>Demographic and Health Survey (DHS). 2006/7.</td>
</tr>
<tr>
<td></td>
<td>Demographic and Health Survey (DHS). 1995–2011 (7 surveys).</td>
</tr>
</tbody>
</table>

*Note: Further information on DHS and LSMS can be found online at [http://www.dhsprogram.com](http://www.dhsprogram.com) and [http://go.worldbank.org/IFS9WG7E00](http://go.worldbank.org/IFS9WG7E00), respectively.*
### Appendix II: Studies using LSMS and DHS in conflict-affected countries

<table>
<thead>
<tr>
<th>LSMS or DHS survey</th>
<th>Academic work</th>
<th>Selected Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>DHS Rwanda 1992 and 2000</td>
<td>De Walque, D. and P. Verwimp. 2010.</td>
<td>• Finds that genocide-related mortality was highest among educated and urban families.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Demonstrates that “low ability” individuals are most likely to be displaced, suffering from the direct effects of conflict and the adjustment costs of displacement.</td>
</tr>
<tr>
<td>LSMS Tajikistan 1999 &amp; 2003</td>
<td>Shemyakina, O. 2006.</td>
<td>• Finds that exposure to conflict has a significant effect on female enrollment and schooling attainment, but no effect on males.</td>
</tr>
<tr>
<td>LSMS Tajikistan 1999 &amp; 2003</td>
<td>Shemyakina, O. 2009.</td>
<td>• Studies the effect of conflict on marriage and reproductive behavior, finding that conflict postpones marriage among women of marriageable age.</td>
</tr>
<tr>
<td>LSMS Tajikistan 1999 &amp; 2003</td>
<td>Justino, P. and O. Shemyakina. 2008.</td>
<td>• Studies the effect of remittances on the labor supply, finding that remittances decrease labor participation rates, especially in conflict areas.</td>
</tr>
<tr>
<td>DHS Cambodia 2000</td>
<td>De Walque, D. 2004.</td>
<td>• Finds that excess mortality during the Khmer Rouge period was especially likely among adult males, especially with urban or educated backgrounds.</td>
</tr>
<tr>
<td>LSMS Kosovo 2000</td>
<td>Bhaumik, G., I. Gang, and M. S. Yun. 2005.</td>
<td>• Studies the relationship between ethnic conflict and economic disparity, showing that, despite an advantageous economic position, Serb rates of poverty were higher than Albanians.</td>
</tr>
<tr>
<td>LSMS Kosovo 2000</td>
<td>Alva, M., E. Murrugarra, and P. Paci. 2002.</td>
<td>• Examines the costs of conflict in education, showing that ethnic tension has harmed Albanian male youth educational attainment.</td>
</tr>
</tbody>
</table>
**Appendix II: Continued**

<table>
<thead>
<tr>
<th>Study</th>
<th>Reference</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSMS Kosovo 2000</td>
<td>Douarin, L., J. Litchfield, and R. Sabates-Wheeler. 2010.</td>
<td>Finds that exposure to conflict predicts the livelihood choices of households, such as the take up of wage labor, reliance on remittances or social assistance, or entrepreneurial activities.</td>
</tr>
<tr>
<td>LSMS Bosnia Herzegovina 2001 – 4 (4 surveys)</td>
<td>E. Kondylis. 2007.</td>
<td>Investigates conflict’s effects on the labor market and finds that “able” workers are more likely to be displaced and unemployed after conflict.</td>
</tr>
<tr>
<td>LSMS Bosnia Herzegovina 2001</td>
<td>E. L. Swee. 2009.</td>
<td>Finds that war intensity, particularly the military draft, adversely affects secondary, but not primary, schooling attainment.</td>
</tr>
<tr>
<td>LSMS Bosnia Herzegovina 2001</td>
<td>Do, Q. T. and L. Iyer. 2009.</td>
<td>Finds no significant differences on mental health from people who experienced different levels of conflict intensity.</td>
</tr>
<tr>
<td>LSMS Nepal 2003/4</td>
<td>M. Hatlebakk. 2007.</td>
<td>Studies the effects of Maoist influence on data collection quality, finding only minor impacts such as the need for approval.</td>
</tr>
<tr>
<td>LSMS Nepal 2003/4 DHS Nepal 1996–2006 (3 surveys)</td>
<td>C. Valente. 2011.</td>
<td>Finds that abductions by Maoists and conflict intensity increased the probability of early marriage but that only abductions by Maoists had a negative effect on school attainment.</td>
</tr>
<tr>
<td>UNFPA- Burundi 2002</td>
<td>Bundervoet, T. 2009.</td>
<td>Finds that older, wealthier and better-educated males were more likely to be killed in the 1993 Burundi massacres.</td>
</tr>
<tr>
<td>36 DHS surveys</td>
<td>Østby, G. 2008.</td>
<td>Finds that communal pressure for land increased the likelihood of killings.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Finds evidence that “horizontal inequalities” that coincide with identity divisions aggravate grievance and promote social cohesion, facilitating mobilization for conflict.</td>
</tr>
</tbody>
</table>

References


The World Bank Research Observer, vol. 31, no. 1 (February 2016)


Promoting Women’s Economic Empowerment: What Works?

Mayra Buvinic´ and Rebecca Furst-Nichols

A review of rigorous evaluations of interventions that seek to empower women economically shows that the same class of interventions has significantly different outcomes depending on the client. Capital alone, as a small cash loan or grant, is not sufficient to grow women-owned subsistence-level firms. However, it can work if it is delivered in-kind to more successful women microentrepreneurs, and it should boost the performance of women’s larger-sized SMEs. Very poor women need a more intensive package of services than do less poor women to break out of subsistence production and grow their businesses. What works for young women does not necessarily work for adult women. Skills training, job search assistance, internships, and wage subsidies increase the employment levels of adult women but do not raise wages. However, similar interventions increase young women’s employability and earnings if social restrictions are not binding. Women who run subsistence-level firms face additional social constraints when compared to similar men, thus explaining the differences in the outcomes of some loans, grants, and training interventions that favor men. Social constraints may also play a role in explaining women’s outcome gains that are short-lasting or emerge with a delay. The good news is that many of the additional constraints that women face can be overcome by simple, inexpensive adjustments in program design that lessen family and social pressures. These include providing capital in-kind or transacted through the privacy of a mobile phone and providing secure savings accounts to nudge women to keep the money in the business rather than to divert it to non-business uses. JEL Codes: J16, J24, L25, L26, M53, O12

The observed disparity between the sexes in productivity and earnings is persistent and pervasive. The value added per worker is between 6 percent and 35 percent lower in female-owned than in male-owned firms; female-managed farms are 20 percent to 80 percent less productive than male-managed farms; in the workplace,
females earn between 20 percent and 80 percent lower average wages than do males, depending on the country (World Bank 2012). This disparity is not because women are inferior entrepreneurs, farmers or wage workers; it is primarily the result of differences in the size of their businesses and farms, in the sectors in which they operate, and in human capital (health and education) and returns to this capital (World Bank 2012, 2013).

A wide range of policies and programs—from strengthening economic rights for women under the law to providing women with greater access to quality child care and financial literacy—can potentially spur women’s economic advancement and reduce gender gaps in economic performance. This article reviews the empirical evidence on the effectiveness of actions that have direct, near-term impacts on women’s economic outcomes while recognizing the importance of complementary investments in women’s human capital and inclusive policies and legal frameworks over the long term.

Short-term interventions that succeed in increasing women’s earnings are not to be underestimated. They benefit women and can have a transformative impact on society by fostering greater investments in child well-being, reduced household poverty, and enhanced aspirations for the next generation of girls and women (World Bank 2012).

The interventions covered in this review include providing women with access to capital (loans and grants) and with savings accounts, business management training, on-the-job and skills training, and job vouchers. These are all increasingly commonplace interventions that are intended to boost the productivity and earnings of self-employed women, women farmers and women wage workers in developing countries. Table 1 summarizes the principal evaluation studies grouped by intervention type. As shown in the table, this review draws on credible evaluations of the effectiveness of different interventions—evidence that is rigorous enough to ensure that the observed outcome was the result of the intervention rather than of selection effects or placement biases. The article updates and benefits from thematic reviews of this evidence commissioned by the Women’s Roadmap Project.1

The main question guiding the review is “what works” to increase women’s productivity and earnings as proxies for the more complex notion of economic empowerment. This question is followed by the question of “for whom,” hypothesizing that the evidence would show differential outcomes for poor or very poor versus non-poor women (defined by income levels, type of occupation, or the nature and size of firms or farms) and for young (between the ages of 15 and 24, or older in post-conflict settings) versus adult women (aged 25 to 64). The gender literature’s emphasis on documenting gender differences has often underplayed the equally important documentation of heterogeneity among women.
Methodological Issues

Interventions were assessed according to the strength of empirical evidence. They were judged as proven when the evidence was robust, with more than one credible evaluation yielding similar results across different settings or circumstances. Promising interventions were backed by positive findings from one credible study or its equivalent. Unproven interventions yielded no evidence of positive effects on productivity and earnings for the categories of women and settings evaluated.

Measuring changes in the productivity and earnings of individual women is straightforward conceptually but much less so practically, especially for women in low-income settings who farm or are self-employed and have highly variable earnings (Rosenzweig 2012; World Bank 2013). In response, many of the studies reviewed used a variety of intermediate outcomes for women working in self-employment and farming, including growth in women’s physical and financial assets, improvements in the stability of women’s earnings, evidence of business creation or start-ups, and increases in women’s leisure time.

A majority of the evaluations reviewed rely on randomized control trials (RCTs) and naturally occurring experiments that yield rigorous evidence of cause-effect relationships but that are not usually very good at providing evidence on the “why,” the channels or mechanisms that explain why the intervention had a particular effect. This review pays special attention to possible channels that mediate observed outcomes. It also examines the timing of outcomes, when they appear and how long-lasting they are, and the cost-effectiveness of interventions.

What the Evidence Says

Access to Capital: Loans and Grants

Poor and very poor women. The experimental evidence (table 1) overall shows that a small infusion of capital alone, as a loan or a grant, is not enough to grow subsistence-level, female-owned firms. Poor women who run subsistence-level enterprises (with average monthly revenues of US $80 to $100 at market exchange rates, profits of approximately $1 per day, and no paid employees) fail to benefit in terms of business profits or growth from micro loans or cash grants (of approximately $100 to $200 on average). This result was observed in Sri Lanka, where cash grants were given (de Mel, McKenzie, and Woodruff 2008, 2009, 2012), and in India, Bosnia and Herzegovina, Mongolia, the Philippines, and Thailand, where the intervention was microloans (Duflo et al. 2013; Augsburg et al. 2012; Attanasio et al. 2011; Karlan and Zinman 2010; Coleman 2006).

In contrast, capital alone increased the business profits of male-owned micro-enterprises in both the Philippines and Sri Lanka – an effect that, in the case of...
Sri Lanka, was still evident five years after the cash was disbursed (Karlan and Appel 2011; de Mel, McKenzie, and Woodruff 2012).

The finding that micro loans or grants have a positive effect on the performance of male-owned firms, but not female-owned firms, may be because subsistence-level female firms operate in sectors that face more severe constraints to growth and because women face more pressures than men to spend some of the cash intended for the business for other purposes, or a combination of both.

Duvendack et al. (2011) argue that the lack of effect of microloans is because many women microloan clients are “necessity” entrepreneurs who turn to self-employment because there are no other jobs available and who would be better off in wage employment. De Mel, McKenzie, and Woodruff (2012) add that women face social and time constraints that limit the types of businesses they operate and the impact of capital alone on business profits and growth.

Alternatively, women working in subsistence-level enterprises may face pressures to use portions of their revenue for other expenses and transfers rather than for business investments. Pressures can have internal or external origins. According to Fafchamps et al. (2014), an internal, present-day orientation rather than external pressures to share the cash seems to guide women’s choices, whereas men show a future orientation or more self-control. Blattman, Fiala, and Martinez (2014) also find young women to be more present-biased than young men, although this orientation did not prevent young Ugandan women from increasing their earnings in response to a large cash grant. An experiment in 26 rural communities in Western Kenya provides ex ante evidence for the external pressure hypothesis. Women invested less in a presumed business venture when investment income was visible compared to when it was hidden, and this tendency was more pronounced when close kin was present. This “kin tax” was applicable only to women; there was no effect for men (Jakiela and Ozier 2012).

Further supporting evidence for external pressures that influence women’s but not men’s diversion of capital from firm to family comes from recent randomized evaluations of combined training and loan provision in Uganda, savings in Chile, and ATM cards in Kenya. Family proximity in Uganda had significant negative effects on the business investment decisions of married women, whereas men benefited from this proximity (Fiala 2014). In Chile, the take-up of personal savings accounts was particularly high among microentrepreneurs (91 percent of whom were women) who were not heads of household and faced significant pressures to share resources with family or community (Kast and Pomeranz 2014). In Kenya, women with low household bargaining power did not use unsecured ATM cards for individual bank accounts, which reduced transaction costs, because of fear they would not be able to resist their husbands’ pressure to obtain some of those savings (Schaner 2014).

By helping protect the privacy of individual choices, mobile money may provide women with sufficient autonomy and independence from social pressures to invest...
cash in the business rather than in the extended family, as Aker et al. (2011) found in a randomized evaluation of a cash transfer program delivered to poor women via mobile phones in Niger. Positive crop outcomes were likely the result of the low cost of using the mobile phone to transfer cash and the greater privacy the mobile phone gave women to elect how to spend the transfer.

The evidence also suggests that capital delivered in-kind (e.g., purchase of livestock or inventory) rather than as cash works for women-owned microenterprises with firms that are above subsistence level in size—that is, those with higher initial profits than other female-owned firms. Large, lasting gains in profit for female-owned microenterprises in urban Ghana were achieved when an in-kind transfer of capital (i.e., purchase of inventory) nudged the recipient women to keep a future business orientation and not divert the money to more immediate uses (Fafchamps et al. 2014).

**Non-poor women.** Access to finance may be enough to increase the growth and earnings of women-owned small firms in the larger, formal small and medium-sized enterprise (SME) segment (with 5 to 19 employees per firm), although rigorous evidence on the impact of access to financial services for SMEs by the firm owner’s sex is still lacking (partly because of the comparatively few women-owned SMEs compared with women-owned microenterprises). However, there is sufficient evidence that in developing economies, capital constraints are one of the main reasons for the suboptimal size of female-owned firms compared with male-owned ones (Klapper and Parker 2011; Sabarwal and Terrell 2008) and, more generally, are a binding constraint to the growth of all small firms in developing countries, both female- and male-owned firms (Ayyagari, Demirci-Kunt, and Maksimovic 2012).

**Young women.** A comparatively large infusion of capital in the form of an unconditional grant without training or other support increased the earnings of poor young adults (aged 16 to 35) in conflict-affected northern Uganda. This effect, still visible four years later, was stronger for young women (who started from lower incomes) than for young men (Blattman, Fiala, and Martinez 2014). It may be that this powerful effect of a large cash windfall on young women’s earnings would extend to adult women running subsistence enterprises, meaning that access to capital alone would grow these adult women’s earnings provided that it was a large enough amount and women were given time to show results.

**Timing of effects.** Panel household survey data for Bangladesh covering a twenty-year period support the notion that time influences the outcomes we observe. These data show a beneficial effect, greater for females than for males, of 20-year cumulative microcredit borrowing on household per capita income and the reduction of extreme poverty (Khandker and Samad 2014). The authors speculate that past credit may affect current credit, leading to more risk taking over time. This could be
particularly relevant for women, who may be more risk averse than men. Similarly, a two-month grace period versus immediate repayment requirements for poor urban women borrowers in Kolkata, India significantly raised long-run (three-year) business profits by encouraging risk taking (Field et al. 2014).

**Capital plus Training**

*Poor and very poor women.* Bandiera et al. (2013a) show that a large capital transfer (most often a cow or other livestock with an asset value of approximately $140), intensive asset-specific training, and regular follow-up visits by an asset specialist and BRAC program officer during a two-year period had a significant, transformative impact on the occupational choices of very poor women in Bangladesh. These women changed occupational choices from casual day labor to self-employment and increased earnings significantly. This positive effect was still evident four years later (Bandiera et al. 2013b). Although all of the women included were very poor, the effect was largest for women who had the highest initial relative earnings.

Two evaluations of very similar pilots for the ultra-poor in rural India, one in West Bengal and one in Andhra Pradesh, yield findings in line with those of Bangladesh, although the simultaneous implementation of a government program for agricultural wage workers confounded the results from the Andhra Pradesh pilot (Banerjee et al. 2011; Morduch, Ravi, and Bauchet 2012; Bandiera et al. 2013a). Bandiera et al. (2013a) reviewed the available evidence for cost-effectiveness and found, for the five studies that had data on costs (including their own), that returns from these capital- and skills-intensive ultra-poor interventions were cost effective when compared to yields from savings accounts.

The exceptions to these positive findings are randomized evaluations of ILO business training programs plus capital in Sri Lanka and Uganda. In urban Sri Lanka, the increased earnings of female business owners who received training and a US$130 one-time grant dissipated after two years (de Mel, McKenzie, and Woodruff 2014). In semi-urban Uganda, the delivery of capital (approximately US$200) plus training increased profits for male microentrepreneurs but not for female microentrepreneurs. For men, the training helped ensure that the short-term results of the loans observed at six months lasted over nine months. The lack of results for women was attributed to family pressures: married women were drained by family demands for cash, whereas married men employed family members, thus benefiting from close family ties. Women who were physically separated from their families benefited from the loan plus training combination (Fiala 2014).

**Savings**

*Poor and very poor women.* The evidence on savings is less rich than that on loans and grants but more consistently shows positive economic outcomes for women.
It also shows that women’s demand for saving accounts is high. A review of nine randomized field experiments in countries covering different regions (with studies in Kenya, Philippines, Nepal and Guatemala reporting data disaggregated by sex) shows that savings are a promising way to improve rural women’s productivity (Knowles 2013). In both Ghana and Rwanda, differences between men and women in the existence of savings accounts explained the largest proportion of the gender wage gap for entrepreneurs (Gamberoni, Heath, and Nix 2013). For Ghana, where panel data were available, savings accounts were especially important to increase the profits of the smallest, subsistence-level women entrepreneurs, shifting business performance from losses to profits.

The recent experimental evidence sheds light on reasons why savings may work especially well for women. In Western Kenya, women with access to savings accounts invested 45 percent more in their businesses and were less prone to sell business assets to address health emergencies (Dupas and Robinson 2009). These effects were sustained in a follow-up survey conducted three years later, showing that savings interventions can have lasting effects (Dupas and Robinson 2013).

In rural Mindanao, Philippines, the decision-making ability and control over assets of less empowered women with below-median reported household decision-making power increased as result of commitment savings, where deposits could only be withdrawn after a savings goal was reached; there was no similar effect for men (Ashraf, Karlan, and Yin 2006, 2010). These savings likely provided women with greater legal and psychological control over funds. Commitment savings were particularly attractive to women who reported a present bias. The authors hypothesize that commitment savings are preferred by Filipino women who are responsible for household finances and recognize the temptation to spend or the risk of spousal control over funds.

In urban Chile, however, there was particularly high take up of free, liquid savings accounts by microentrepreneurs (91 percent of them female) who were more exposed to family pressures than others. Liquid savings reduced reliance on short-term debt, improved consumption smoothing in the face of economic shock (indicating an advantage of liquid over commitment savings) and still protected resources from being shared with others (Kast and Pomeranz 2014).

Women’s demand for savings accounts is suggested by the data. In urban Ghana, more than half of women microentrepreneurs use a *susu* collector (a vehicle for informal mobile savings), even when they must pay for the service (Fafchamps et al. 2014). In Western Kenya, women make use of savings accounts through village cooperatives far more than men do (Dupas and Robinson 2009). In Guatemala, female microfinance borrowers are significantly more likely than their male counterparts to open a commercial savings account (Atkinson et al. 2010). In the Philippines, women are more likely to take up commitment savings, although the difference compared with men is not statistically significant (Ashraf, Karlan, and Yin 2006).
In Nepal, a majority (84 percent) of females in 19 urban slums opened simple savings accounts when offered access (Prina 2013).

**Business Training**

**Poor and very poor women.** Both microfinance organizations and banks usually offer business management training courses of varying length and quality as well as varying depth and breadth of business topics covered, which complicates the task of measuring their impacts. A systematic analysis by Woodruff and McKenzie (2013) of 20 evaluations of business training interventions that included female trainees and a complementary review by Bandiera et al. (2013a) of field experiments that combine business training and capital transfers consider this training heterogeneity and conclude that business training improves business practices for women micro-enterprise owners but has few measurable effects on business survivorship or profits. Evaluations across a wide spectrum of countries, including Bosnia and Herzegovina (Augsburg et al. 2012), Pakistan (Gine and Mansuri 2014), Peru (Valdivia 2011), Sri Lanka (de Mel, McKenzie, and Woodruff 2012), and Tanzania (Oppdal Berge, Bjorvatn, and Tungodden 2011), support these conclusions.

The lack of positive training outcomes is not restricted to women but is more consistently observed for women than for men and is attributed to women’s lower-quality businesses and more binding external constraints, including greater household duties and limited say in household decisions (Gine and Mansuri 2014; Oppdal Berge, Bjorvatn, and Tungodden 2011). In addition, attrition rates in business training programs are systematically higher for female than for male trainees, suggesting that these programs may not be designed to respond to the specific constraints women in business face (Valdivia 2013; Woodruff and McKenzie 2013).

Business training seems to have more of an impact on revenues and profits for business start-ups. Training helped women who were out of the labor force start enterprises more quickly in Sri Lanka. Getting these women to start businesses was easier than getting these businesses to grow (de Mel, McKenzie, and Woodruff 2014). Additionally, business training increases microloan repayment rates (Woodruff and McKenzie 2013).

**Non-poor women.** Woodruff and McKenzie (2013) suggest that quality business training on its own may work better for large versus small firms. In the case of large firms, a small percentage increase in profits may be enough to justify the costs of business training, especially if the impacts last, because the costs of training often increase less than proportionally with the size of firms undergoing training. Xu and Zia (2012) similarly suggest that business training has better results with women-owned SMEs when compared with women-owned microenterprises. They advise
targeting motivated, self-selected women entrepreneurs to increase the impact of training programs.

**Timing of effects.** Two- to three-day training programs are very common, though short courses are likely to have less impact than longer ones. Supporting the notion that high-quality training of reasonable duration can make a difference, a six-week, fully subsidized, high-quality basic business education for poor female microentrepreneurs in rural Mexico had noticeable effects, largely by improving accounting practices (Calderon, Cunha, and de Giorgi 2013). Similarly, a three-month intensive training delivered by specialized professionals increased business sales of women microentrepreneurs in Peru. Although in the short term only training plus customized technical assistance to the firm showed an effect on sales, women who received training without technical assistance were able to catch up and show increased sales approximately two years after the training was first offered. Women had adopted the recommended business practices and made business adjustments, albeit at a slower pace than those who also received technical assistance. The business training appeared to be cost effective, unlike the customized advice to the firm, which cost twice as much as the training but whose effects were no longer visible in the second year (Valdivia 2011, 2013).

**Skills Training, On-the-Job Training and Job Vouchers**

**Poor women.** Evaluations of large-scale government programs in Latin America (three in Argentina and one in Mexico) that provided employment vouchers (around $100–$150 per month) that function as wage subsidies and/or skills training to the unemployed poor in times of economic hardship find high uptake by females and positive impacts on their employment, but less clear effects on their wages (Galasso, Ravallion, and Salvia 2004; Jalan and Ravallion 2003; Galasso and Ravallion 2004; Revenga, Riboud, and Tan 1994).

Evaluations of retraining and wage subsidy interventions in four transition economies (Poland, Romania, Russia, and Slovakia) yield mixed effects: positive effects for women’s employment probabilities in Poland and Romania, no effects in Russia and mixed effects in Slovakia (Kluve, Lehmann, and Schmidt 1998; Benus et al. 2005; Lubyova and Ours 1999). Again, there is less support for the effect of these programs on wages.

Finally, a retraining program in China also has mixed findings: a negative impact on employment probability in one city and a positive impact in another city, with no differences between the sexes and no discernible effects on wages in either location (Bidani, Goh, and O’Leary 2002). The program was better implemented in the second city, which may explain the observed city difference in employment outcomes (Todd 2013).

Buvinic and Furst-Nichols 67
Todd’s (2013) review of the evaluation evidence argues that these programs helped workers find jobs more quickly, but at the expense of lower wages than workers may have obtained from a longer job search on their own. This analysis and a review on the same topic by Betcherman, Olivas, and Dar (2004) conclude that these large government programs tend to increase employment but not wages, and they raise the issue of possible negative ‘spillover’ effects on the jobs and wages of people who do not participate as well as the possibility that publicly provided training could crowd out training provided by private sector firms.

Young women. In contrast to the above findings, youth labor training programs in Latin America that provide classroom training, internships and stipends for child care have documented success in promoting young women’s employment. Young women have flocked to these programs and have generally outperformed young men in obtaining results, as evidence has shown for programs in Peru (Nopo, Robles, and Saavedra 2007), Colombia (Attanasio, Kugler, and Meghir 2011), and Argentina (Galasso, Ravallion, and Salvia 2004). The exception is a youth employment program in the Dominican Republic, which had a stronger impact on young men’s employment in formal sector jobs compared with young women’s but did reduce teen pregnancies among female trainees and increase their social and personal competencies (Ibarrarán et al. 2012).

These programs’ overall success is related to the soft skills they offer in addition to technical skills, the quality of on-the-job training, and the commitment of firms to hire trainees. ‘Soft skills’ are personality traits, motivations and preferences that are valued in the labor market, and recent evidence shows that they matter and should have an important place as part of job training programs (Heckman and Kautz 2012). The comparatively greater success of these programs among young women versus young men could be because of systematic gender differences in performance by female trainees, because of firms’ preference for hiring female over male workers, or both (Katz 2013). However, young women’s outperformance could result from the fact that in the control groups, young men have an easier time finding jobs, whereas young women stagnate.

Timing of effects. Recent evidence from Liberia, Uganda, and India replicates the positive outcomes of well-designed skills training programs for young women that are sustained in follow up measures, even in contexts where wage work is scarce. In Liberia, a demand-driven skills training program increased young women’s employment and earnings immediately and in a follow-up measure six months after the classroom training ended (Adoho et al. 2014). Gains were particularly large for young women using business skills learned in self-employment. In Uganda, life skills and vocational training plus adolescent girls clubs increased the earnings of adolescent girls and delayed marriage, reduced fertility and decreased risky health
behaviors. The authors attribute the successful outcomes evident two years after the program to the combined effect of health and economic empowerment interventions (Bandiera et al. 2014). In India, a subsidized six-month long vocational training program in stitching and tailoring for young women residing in poor slums of New Delhi increased employment, self-employment and earnings substantially over controls. Positive effects were sustained 18 months after the training ended and were attributed to an increase in skills rather than an increase in self-confidence or the positive effect of having a training certificate (Maitra and Mani 2014). Analyses showed that the programs in Liberia, Uganda, and India were all cost effective.

In Jordan, however, the combination of vouchers and training increased the employment of young women graduates in typically female occupations in the short term, but this effect disappeared in a follow-up measure 14 months later when the vouchers had expired. The program was not able to overcome social constraints that prevent firms in Jordan from hiring young women. The soft skills training component, however, led to sustained improvement in life outlook and reduced depression in young women (Groh et al. 2012).

Skills training programs can suffer from high female attrition rates, similar to business training programs. High attrition may be more pronounced in programs that do not pair skills training with job internships, vouchers, or stipends. More than 40 percent of female trainees did not complete the six-month training program on stitching and tailoring in New Delhi; significant barriers included the distance to the training center and lack of available child care support (Maitra and Mani 2014). One-third of participants in a three-month long apprenticeship program for poor vulnerable youth (orphans or school dropouts) aged 15 to 24 in Malawi dropped out. Female participants dropped out much more often because of external constraints, whereas their male counterparts did so to take advantage of unrelated job opportunities. The training was more costly to women, who had less access to financing, used more of their personal savings, and were treated worse than men during the apprenticeship (Cho et al. 2013).

The target population in most of these studies is women and young women among the unemployed poor or first job seekers. Unfortunately, the studies do not disaggregate findings within different poverty levels or groups.

Discussion: What Does Work

Heterogeneity Affects Outcomes

It is clear from the review of the evidence that for whom (the client) matters when assessing what works to promote women’s economic empowerment. Few evaluated interventions work for all women across different socioeconomic and age groups.
The more rigorous evidence, which better describes the characteristics of beneficiaries, often shows that the same class of interventions has significantly different outcomes depending on the client or beneficiary. This is the case with small infusions of capital, such as loans or grants, which work for the non-poor but not for the very poor if delivered as a stand-alone intervention. Capital alone, as a small loan or a grant, is not enough to grow women-owned subsistence-level firms. However, it can work if it is delivered in-kind (e.g., inventory) to more successful women microentrepreneurs with somewhat larger businesses, and it should boost the performance of women’s larger sized firms in the SME sector regardless of the way it is delivered.

Something similar occurs with business management training. Business training improves business practices but has few measurable effects on the growth of women-owned subsistence-level firms. Its impact may be improved by targeting women running larger-sized firms as well as by increasing the quality and duration of the training. To see effects, sufficient time should be allowed before outcomes are measured.

Further, the evidence shows that very poor women need a more intensive package of services than do less poor women (and poor men) to break out of subsistence production – on and off the farm, in agriculture and entrepreneurship – and to reap the benefits of higher productivity and earnings for themselves and their families. These interventions are expensive but can be cost effective. To move into profitable work, very poor women subsistence producers and entrepreneurs require a sufficiently large capital transfer paired with business training and follow-up technical visits. There is no silver bullet to grow the productivity of subsistence producers. This situation is particularly challenging because the delivery of bundled services is more expensive and requires more implementation capacity, and the very poor often reside in regions with few resources, frail institutions, and weak delivery systems.

What works for young women does not necessarily work for adult women. Skills training, job search assistance, internships, and wage subsidies introduced to cope with economic shocks increase the employment levels of adult women but do not raise wages. However, similar skills training combined with on-the-job training or wage subsidies increase young women’s employability and earnings if social restrictions against hiring young women are not binding. Different levels of constraints, labor market experience and opportunities between young and adult women are likely responsible for the noticeably better outcomes of these job programs among young women. Young women most likely face fewer social constraints than do their adult counterparts. They may also feel less pressure to take any job and may be more selective in training choices. Nevertheless, the better-performing programs with young women also offer a combination of services, including stipends or vouchers, which increase their costs, but these can still be cost effective.
Interventions proven to increase young women’s economic opportunities are especially attractive given the large potential benefits for those just entering the labor market, including potentially delaying early childbearing.

**Women Face Additional Constraints**

It is also clear that pre-existing gender differentials and inequalities influence the outcome of interventions. Women running subsistence-level firms face additional social constraints to improving their productivity and earnings when compared to similar men, thus explaining differences in the outcomes of some (loan and grant) interventions favoring men. Growing evidence strongly suggests that women running subsistence-level firms face external pressures to divest some of the windfall cash from loans or grants to relatives or household expenses. Thus, both in-kind capital and capital transacted through the privacy of a mobile phone have better outcomes because they nudge women to keep the money invested in the business. Men do not appear to face these same pressures.

This same mechanism (i.e., external pressures) seems to account for the finding that savings accounts are a good vehicle for increasing women’s business earnings. Saving might elicit a greater sense of personal ownership among women compared with cash from a loan or a grant, which would make them less likely to divert savings intended for a business investment to household expenditures or to relatives. Further, savings may be especially beneficial to more risk-averse women, who may not be as comfortable as less risk-averse women with the repayment and interest requirements of a loan. In addition, the evidence shows that savings may particularly benefit both the smallest, subsistence-level women entrepreneurs and women with little household decision-making power, which makes savings an intervention that can apply to women across the board, regardless of their poverty condition and the social restrictions they face.

Women’s comparative disadvantage is also evident in the evaluations of training programs. The low uptake and high attrition rates in many of the business and skills training programs reviewed suggest that attendance costs (both time and money) are higher for women, who face more social constraints to participation, than for men. High attrition rates affect these programs’ effectiveness and cost effectiveness as well as the validity of program evaluations (by reducing sample size and, therefore, the power of statistical tests). All indications are that many of these programs are not designed with women clients’ constraints in mind.

**Clever Design Adjustments are Available**

The good news is that many of women’s additional constraints can be overcome by simple, inexpensive adjustments in program design. This review has highlighted
simple, clever design features that help women overcome family and social pressures as well as the mobility constraints that impede women from growing their earnings in poorer and/or more socially conservative settings. A telling example is to provide capital in-kind (e.g., by purchasing inventory or a physical asset) rather than in cash to nudge women to keep the money in the business rather than to divert it for household use or pass it on to relatives.

Many of these design features give women a measure of autonomy in an otherwise restricted environment where they are not free to make work-related choices or travel alone. Perhaps most notable among them is the use of mobile money services to conduct financial transactions in private, receive reminders to save and obtain information on prices and jobs in real time without having to travel long distances. Financial services delivered through mobile phones can be less costly and may be more effective than other delivery mechanisms in increasing women’s business investments. However, security must be built into financial products that use new technologies to reduce transaction costs. Unsecured accounts may not be attractive to women who may fear capture by more powerful husbands or relatives. Secure accounts and commitment savings may be preferred by women with family pressures or self-control problems. However, because savings liquidity may be useful to cope with economic or other shocks, the design of savings should strive to balance both the security and liquidity of accounts, and, in some cases, women may find uses for both types of accounts. Minor design adjustments in business and skills training programs and in class schedules and locations could go a long way toward increasing women’s uptake, retention, and training effectiveness. More generally, minor adjustments to the design of financial services and business and skills training to better accommodate the needs and demands of women clients could improve performance and outcomes considerably.

Market research could usefully test savings and loan products and training offerings demanded by different categories of women clients (by occupation and income levels, for instance), contributing to the design of products that are more appropriately tailored to women entrepreneurs and farmers.

**Female Autonomy Helps**

Interventions seem to work best when women micro and small producers and entrepreneurs have economic independence or autonomy and few social restrictions. Female autonomy means that women have physical mobility and can make independent choices, both of which are requirements to succeed in business. Autonomy allows women to seek and benefit from productive services and resources directly, increasing the likelihood that they will be able to switch from low- to high-value production. Autonomy is either measured directly or is indicated by the fact...
that women are unmarried, divorced or widowed heads of household or sole income-earners and, therefore, do not have a male partner who can control their choices. Sole-income earners were a significant proportion of the ultra-poor women in Bangladesh who transformed their occupations as result of the capital and training provided by BRAC. Female headship was a characteristic trait of women farmers in Malawi and urban entrepreneurs in Uganda who successfully crossed over from lower-value to high-value production. (Dimova and Gang 2013; Campos et al. 2013).

An alternative explanation of the above evidence is that these women entrepreneurs and farmers do not necessarily lack autonomy, but they access fewer productive services because their provision has traditionally been biased toward male clients, especially when there is an adult male present in the household.

Autonomy or its absence seems to be much less of a concern for adult and young women in wage employment, however. This may be partly because women have higher status and are more independent in countries where the wage sector is more prominent and employs more women. However, it may also be that wage employment in itself provides women with a measure of autonomy from the family, at least within the confines of the workplace.

**Time Can Influence Observed Outcomes**

Project impacts are not all necessarily linear or monotonic, and evaluation timetables may not allow sufficient time for results to occur or to understand if effects are long-lasting or disappear with time.

Evaluations that include repeated measures sometimes show positive employment and income effects for women that disappear with time, as in the case of the combination of training and a cash grant to Sri Lankan female business owners and vocational training and job vouchers to Jordanian young women. Other times, they show effects that emerge with a delay, as in the case of two decades of microenterprise lending in Bangladesh and business training for microentrepreneurs in Peru. Gender norms seem to play a role in influencing outcomes over time. Social restrictions specific to women can override project gains. However, project gains can also help to change gender norms, causing positive outcomes to emerge with a delay as, for instance, when interventions drive changes in men’s initial resistance to women working outside the home or in women’s own avoidance of business risks.

Rigorous evaluations increasingly include repeated measures to assess the lasting impact of programs. This practice is to be encouraged, especially with the type of short-term economic empowerment interventions for women that are the focus of this article.
Knowledge Gaps

There is very little knowledge on the characteristics and determinants of the success of working women—successful entrepreneurs and farmers who have “crossed over” to high-value production, young women who have successfully transitioned from school to good jobs in the labor market, and wage workers who have moved up to higher-paid jobs in factories or firms or into non-traditionally female occupations. This gap in knowledge is not surprising because there are comparatively few women to study. However, a systematic study of the trajectories of success among working women across sectors should yield valuable information for program and policy design.

The results of this research could be useful to help identify the profiles of those women entrepreneurs, farmers or wage workers with the firms or farms that are most primed for growth or most ready to succeed in non-traditional occupations. The ability of programs to recruit and select adult and young women who are most ready to succeed in the sector would significantly increase their effectiveness.

This review of the evidence has been limited, by definition, to literature that has been evaluated and published. Evaluations that yield no significant findings are seldom published, so this review contains less information than it should on lessons from failed programs. More importantly, it does not cover promising innovations for which there are no good evaluation data.

Potentially effective interventions that are in need of evaluation data include job information for young women, firm certification programs for wage workers, and non-traditionally female skills training for women wage workers and entrepreneurs. Further evaluation is also needed for business associations and networks for entrepreneurs as well as mentors and role models for young women. Both of the latter programs are well known and frequently implemented, but we found little rigorous information on their impact on women’s productivity and earnings in developing countries. There is only one rigorous study of SEWA bank clients in India, which suggests that the presence of peers positively influences women’s business behavior and changes aspirations, especially among those exposed to more restrictive social norms (Field et al. 2014). These results suggest that business networks and mentors are useful and potentially transformative, but more data are needed about their workings and their labor market impacts.

Finally, as this review has shown, there is a growing body of rigorous evidence on what works to empower women economically, but there is very little evidence on the financial and economic costs and benefits and even less evidence on the sustainability of interventions, both in terms of the lasting impacts of programs and the institutional and political dimensions of sustainability. The gaps in knowledge are large and cut across the reviewed interventions.
Table 1. Principal Evaluation Studies by Intervention Type

<table>
<thead>
<tr>
<th>Study</th>
<th>Features</th>
<th>Sample</th>
<th>Intervention</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Access to Capital: Loans and Grants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Aker et al. (2011) | Mobile cash transfer in rural Niger | Recipient households in 96 villages | Use of mobile phones to distribute unconditional cash transfers in targeted villages | -Mobiles saved time and transport costs.  
-Purchased on average .86 more types of food and non-food items.  
-Had more non-durable assets.  
-Grew .36 more types of crops and more female-produced cash crops. |
| Attanasio et al. (2011) | Group and individual lending in rural Mongolia | 1,148 adult women from 40 communities | Opening of microfinance branches and comparing group liability loans, individual loans or having no branch | -Business ownership increased, but profits and income did not.  
-Higher educated increased business activity in services, whereas lower educated did so in agriculture.  
-Loans did not benefit poorest among marginal clients.  
-Without much human capital, it is difficult to put loans to good use. |
| Augsburg et al. (2012) | Microfinance and education in Bosnia and Herzegovina | 995 poor women who were originally rejected by microfinance institution | Microfinance loans | -Borrowers started and expanded their small businesses.  
-Women increased earnings.  
-Young women showed greater present bias than young men. |
<table>
<thead>
<tr>
<th>Study</th>
<th>Features</th>
<th>Sample</th>
<th>Intervention</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Blattman, Fiala, and Martinez (2014) | Skilled self-employment in rural Uganda | 265 groups of average 22 members each, aged 16 to 35       | Large, unconditional cash grant in groups (equivalent of $382/member) | - After four years, business assets increased by 57%, work hours by 17%, and earnings by 38% compared to controls.  
- Half practiced a skilled trade.  
- Many formalized enterprises and hired labor.  
- Young women’s average incomes were 73% higher than controls compared with young men’s, which were 29% higher, though women started poorer.  
- Women’s work and earnings stagnated without the program. |
| Coleman (2006)             | Microfinance in rural northeast Thailand | 445 households, mostly small-scale farmers                  | Credit Some savings, which may be used to provide loans             | - Positive returns to capital for men’s businesses; mixed impact on women’s businesses.  
- Women invest very little of smaller grants in business but as much, if not more, than men of the larger grant.  
- Women do not experience permanent increases in business income from grants, whereas men do. |
<table>
<thead>
<tr>
<th>Study</th>
<th>Category and Context</th>
<th>Sample Size</th>
<th>Impact Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>de Mel, McKenzie, and Woodruff (2008)</td>
<td>Returns to capital in urban Sri Lanka</td>
<td>385 microentrepreneurs (50% women)</td>
<td>Grants of $100 and $200 in cash or in-kind - No significant impact on earnings for women, whereas impacts are large and significant for men. - Returns vary with entrepreneurial ability and household wealth but not with risk aversion or uncertainty.</td>
</tr>
<tr>
<td>de Mel, McKenzie, and Woodruff (2009)</td>
<td>Gender, credit constraints, and microenterprise returns in peri-urban Sri Lanka</td>
<td>405 low-capital microentrepreneurs (50% women)</td>
<td>Grants of $100 and $200 in cash or in-kind - No significant impact on earnings for enterprises owned by women. - Treatment impacts significant for larger, male-owned enterprises.</td>
</tr>
<tr>
<td>de Mel, McKenzie, and Woodruff (2012)</td>
<td>Long-term effects of one-time capital transfers to microenterprises in urban Sri Lanka</td>
<td>408 microentrepreneurs (50% women)</td>
<td>Grants of $100 or $200 in cash or in-kind; five year follow-up measure - Ten percentage point higher survival rates and $8–$12 higher monthly profits for male-owned businesses. - No short- or long-term impact of capital on female-owned businesses.</td>
</tr>
<tr>
<td>Study</td>
<td>Features</td>
<td>Sample</td>
<td>Intervention</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
</tbody>
</table>
| Duflo et al. (2013)         | Microfinance in urban India             | 52 slums selected to open a branch of a microfinance institution that lends only to women | Microloans only, majority group loans           | -Consumption no different and average businesses no more profitable, although an increase in profits at high end.  
-Men showed more future orientation or self-control.  
-Women had present-day orientation and external pressures to share their cash. |
| Fafchamps et al. (2014)     | Microenterprise growth in urban Ghana   | 800 microentrepreneurs                                                | Cash or in-kind grants to microenterprises       | -For women in subsistence enterprise, no gain in profits from either treatment.  
-For women with larger businesses, in-kind grants increase profits, whereas cash grants do not.  
-For men, cash also has lower impact, but differences between cash and in-kind grants are less robust than for women. |
| Jakiela and Ozier (2012)    | Familial connections (kin) and investment in rural Kenya | 2,145 participants in 26 communities                                  | Game where subjects could vary amount to be invested in presumed business; conditions varied degree of privacy of investment choice | -Women invest 22.1% less when investment income is observable than when it is hidden.  
-Effect is pronounced when relatives are present (equivalent of 4% “kin tax”).  
-No similar tendency to hide income among men. |
<table>
<thead>
<tr>
<th>Study</th>
<th>Setting</th>
<th>Target Group</th>
<th>Credit Access Details</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karlan and Zinman (2010)</td>
<td>Credit access in urban South Africa</td>
<td>Poor men and women wage workers who would have otherwise been rejected for loans</td>
<td>Individual credit with median loan size of $127</td>
<td>Positive impact of access to credit on clients’ retention of jobs (loans likely helped clients smooth or avoid shocks that prevent them from getting to work).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive impact of access to credit on household incomes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Loans increased household food consumption.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No significant difference in impact of credit assigned to men and women.</td>
</tr>
<tr>
<td>Khandker and Samad (2014)</td>
<td>Microcredit in rural Bangladesh</td>
<td>2,600 households in 87 villages (14.7% attrition rate over total years)</td>
<td>20 year panel survey measuring impact of cumulative microcredit borrowing</td>
<td>Increased household non-land assets (by 0.25%).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Increased income and expenditure.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Increased labor supply of men (by 0.33%) and women (by 0.46%).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Reduced extreme poverty, especially for female borrowers, by 5 percentage points.</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Study</th>
<th>Features</th>
<th>Sample</th>
<th>Intervention</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Capital plus Training</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Bandiera et al.      | Basic entrepreneurship    | 1,400 communities of ultra-poor     | Large capital transfer (value about $140), intensive asset-specific training, and regular follow up technical visits | -After two years, very poor women changed occupational choices from casual day labor to self-employment and increased earnings by 34%.  
-Effect was largest for women who had highest relative earnings at the start.  
-After four years, women’s earnings were up 38% relative to baseline, whereas controls remained unchanged.  
Reduction in casual wage labor was twice that of the second-year data.  
-Hours spent in self-employment were lower after four years, likely due to increased efficiency in livestock rearing activities because earnings per hour were 15% higher than at baseline.  
-Significant improvements in entrepreneurial skills.  
-72% increase in likelihood of engaging in self-employment.  
-Gains in employment did not come at expense of school enrollment.  
-26% lower fertility rates in treatment and 6.9 percentage points less likely to be married after two years. |
<p>| (2013b)              | in rural Bangladesh       | women across 40 regions              |                                                                                                   |                                                                                                                                                                                                          |
| Bandiera et al.      | Women’s empowerment       | 4,888 adolescent girls in 100       | Vocational and life skills training                                                               |                                                                                                                                                                                                          |
| (2014)               | in rural and urban Uganda| treatment communities               |                                                                                                   |                                                                                                                                                                                                          |</p>
<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Sample Size</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banerjee et al. (2011)</td>
<td>Targeting the hard-core poor in rural India</td>
<td>1,000 households</td>
<td>Large asset transfer with intensive training on managing the asset</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- After 18 months, average monthly per capita earnings of households who received the asset increased by 21%.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- 15% increase in household consumption.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Positive impacts on household assets and emotional well-being.</td>
</tr>
<tr>
<td>de Mel, McKenzie, and Woodruff (2014)</td>
<td>Business training and cash grants to microenterprises in urban Sri Lanka</td>
<td>624 women subsistence level business owners and 628 women potential business owners</td>
<td>ILO business training plus cash grant of $130 conditional on training completion</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Training changes women's business practices but has no impact on business profits, sales, or capital stock.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Grant plus training increases business profitability in the first eight months, but impact dissipates in the second year.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- For potential startups, business training speeds entry (without changing longer-term ownership rates) and increases profitability.</td>
</tr>
<tr>
<td>Study</td>
<td>Features</td>
<td>Sample</td>
<td>Intervention</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------------------</td>
<td>--------------------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Fiala (2014)</td>
<td>Loans, grants and training in semi-urban Uganda</td>
<td>1,550 microentrepreneurs</td>
<td>Random selection to receive loans, cash grants, business skills training or a combination</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morduch, Ravi,</td>
<td>Anti-poverty program</td>
<td>3,485 ultra-poor women</td>
<td>Large asset transfer with intensive training on managing the asset</td>
</tr>
<tr>
<td>and Bauchet</td>
<td>in rural India</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2012)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Savings</strong></td>
<td>Female empowerment in the rural Philippines</td>
<td>4,000 adult women microfinance clients</td>
<td>Individual commitment savings</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------</td>
<td>----------------------------------------</td>
<td>-------------------------------</td>
</tr>
</tbody>
</table>
| **Ashraf, Karlan, and Yin. (2006, 2010)** | - Significant increase in average savings (of 81% over a year) compared to controls  
- Present-biased women were 15.8 percentage points more likely to take commitment savings.  
- Less empowered women showed largest gain in decision making ability through savings |
| **Atkinson et al. (2010)** | - Prompting for savings at the time of loan payments doubles savings, whereas suggesting a savings deposit equal to 10% of the loan repayment causes savings to double again.  
- Women are significantly more likely to take up the offer of a savings account.  
- Women are more likely to take up commitment savings.  
- Women’s accumulated net savings are significantly lower than men’s overall. | 1,375 borrowers from 20 microfinance branches of Guatemala’s largest public sector bank | Introduction of new commercial savings products to existing borrowers, with no financial incentives or penalties |

*Continued*
<table>
<thead>
<tr>
<th>Study</th>
<th>Features</th>
<th>Sample</th>
<th>Intervention</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dupas and Robinson</td>
<td>Savings constraints in rural Kenya</td>
<td>185 microentrepreneurs</td>
<td>Individual commitment savings with interest-free account and high withdrawal fees</td>
<td>- Positive impact of savings on business investment among women (45% increase). &lt;br&gt; - Increase in women’s private expenditures (37% to 40% higher). &lt;br&gt; - Some impact on making women less vulnerable to health shocks; women were less prone to sell businesses to address health emergencies. &lt;br&gt; - No effect for men. &lt;br&gt; - Positive effects of the savings sustained in a follow-up survey three years later. &lt;br&gt; - Women made use of savings accounts far more than men did.</td>
</tr>
</tbody>
</table>
Kast and Pomeranz (2014) Access to formal savings in urban Chile 3,500 low-income microfinance institution clients (90% women) Providing free and easy access to a formal savings account

- Savings and credit are used as substitute inputs for consumption smoothing. Reducing barriers to savings reduces reliance on debt.
- When cost of savings was reduced, demand for short-term credit decreases and overall consumption smoothing increases.
- Consumption cutbacks associated with a negative income shock reduced by 44%.
- Substantial improvements in subjective well-being.
- Take-up is high among participants who are not head of the household, who have conflicts with partners over money, and who are expected to loan to social network.
<table>
<thead>
<tr>
<th>Study</th>
<th>Features</th>
<th>Sample</th>
<th>Intervention</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Prina (2013)| Savings accounts in urban Nepal   | 1,118 female household heads aged 18–55 | Access to a fully liquid bank account with no opening, maintenance, or withdrawal fees | - After one year, access to savings accounts increased monetary assets by more than 50%.  
- Total assets (monetary and non-monetary) increased by 16%.  
- Positive effect on monetary assets was strongest for poorer households and those not linked to formal banking institutions prior to the intervention.  
- Lower transaction costs due to proximity to the bank and lack of fees may have improved take-up and usage.  
- Saving in accounts rather than cash reduces temptation to spend immediately. |
<table>
<thead>
<tr>
<th>Study</th>
<th>Description</th>
<th>Methodology</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schaner (2014)</td>
<td>Transaction costs, bargaining power, and savings account use in rural Kenya</td>
<td>1,114 newly opened bank accounts owned by 749 married couples</td>
<td>Provision of ATM cards free of charge, when without ATM cards, bank accounts had $.78 withdrawal fee. When provisioned, overall account use among men and married couples’ joint accounts increased by 70% (4.7 percentage points). ATM cards had no impact on female-owned individual accounts. Both men and women with low levels of bargaining power responded negatively to ATM, whereas both men and women with high levels of bargaining power responded positively (controlled for time inconsistency and financial literacy). ATM cards significantly increased the share of individuals reporting that both spouses make joint spending and savings decisions.</td>
</tr>
</tbody>
</table>

**Business Training**

<p>| Calderon, Cunha, and de Giorgi (2013) | Business literacy in rural Mexico | 875 female entrepreneurs (724 in follow-up) of small firms in poor state in Mexico; majority of sample with below secondary education | 6 weeks of business literacy classes with two four-hour meetings per week; total of 48 hours | - 24% increase in profits and 20% increase in revenues, largely through improved accounting practices. - 50% of non-attriting businesses had closed by the time of second follow-up survey 28 months after training. |</p>
<table>
<thead>
<tr>
<th>Study</th>
<th>Features</th>
<th>Sample</th>
<th>Intervention</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gine and Mansuri (2014)</td>
<td>Constraints to entrepreneurship in rural Pakistan</td>
<td>4,100 microfinance clients in groups</td>
<td>Eight full-time days of business training and a lottery to access business loans of USD $1,700 (7 times average loan size)</td>
<td>-Improved business knowledge of both male and female entrepreneurs, but little effect on business outcomes for women (some effects for men). -Only impact for women is improved business knowledge. -Reduced business failure rates, improved business practices, and increased household expenditures by $40/year, but effects concentrated with male clients.</td>
</tr>
</tbody>
</table>
Field et al. (2014) Peer support and entrepreneurship in urban India
636 customers age 18–50 of India’s largest women’s bank (SEWA)
Two days of business counseling and assistance in identifying a medium-term financial goal, with a random subsample invited to attend training with a friend

- Women who attended training alone used loans for home repair, whereas those invited with a friend used loans for business purposes.
- After four months, those who came with a friend reported differences in business behavior (higher volume of business and more stated plans to increase revenues), whereas those who came alone experience no change.
- Those invited with a friend report significantly higher household income and expenditures and were less likely to be housewives.
- Impacts of peer training on business loans and labor supply are concentrated among women in social castes/religions with more restrictive social norms.
<table>
<thead>
<tr>
<th>Study</th>
<th>Features</th>
<th>Sample</th>
<th>Intervention</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Oppedal Berge, Bjorvatn, and Tungodden (2011) | Human and financial capital for microenterprise development in Tanzania | 644 microfinance clients | Business training and a lottery to access business loans | -Improved business knowledge of both male and female entrepreneurs, but increased profits of male-owned firms only.  
-Lack of positive outcomes for women was attributed to their more binding external constraints (i.e., their greater household duties and limited say in household decisions) and their presumed reluctance to compete. |
| Valdivia (2011) | Training or technical assistance in urban Peru | 1,983 female microentrepreneurs | 36 3-hour business training sessions over 12 weeks; half additionally offered individualized technical assistance | -The combination of training and tailored business advice resulted in an average 19% increase in business sales.  
-The sales of women who were exposed only to business training did not grow significantly compared to controls.  
-Having an expert visit the firm to offer customized advice cost twice as much as the training alone. |
<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Participants</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valdivia (2013)</td>
<td>Business training plus in urban Peru</td>
<td>1600 female microentrepreneurs</td>
<td>Same as above – follow-up 2 year measures after training and TA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Training plus tailored TA reported increased sales 7 to 10 months after the intervention.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Those that received training only ‘catch up’ 12 to 15 months later and increase business sales of more than 15% approximately 2 years after end training, especially larger firms.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Take up rate slightly above 50%. Travel time and child care needs affected retention.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Extra cost of TA not cost effective</td>
</tr>
<tr>
<td>Adoho et al. (2014)</td>
<td>Adolescent girls’ employment program in urban Liberia</td>
<td>1,989 young women aged 16–27 with basic literacy and numeracy skills, not enrolled in school, and residing in or near Monrovia</td>
<td>Livelihood and life skills training and assistance with job placement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Increased employment by 47% and earnings by 80%.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Positive effects on empowerment measures, including access to money, self-confidence, and anxiety about future.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- No net impact on fertility or sexual behavior.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Cost-benefit analysis: Budget cost of EPAG business development training is equivalent to three years’ increased income.</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Study</th>
<th>Features</th>
<th>Sample</th>
<th>Intervention</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attanasio, Kugler, and Meghir (2011)</td>
<td>Vocational training in urban Colombia</td>
<td>Unemployed young people placed in lowest two deciles of income distribution</td>
<td>3 months in-classroom training and 3 months of on-the-job training; trainees receive daily stipend, including extra for child care</td>
<td>-11% increase in likelihood to be in paid employment. -Wage and salary earnings increase by 18%. -Controlling for training institution fixed effects and pre-treatment characteristics: -Trainees work an average of 1.1 more days/month and 2.5 more hours/week than women in the control group.</td>
</tr>
<tr>
<td>Benus et al. (2005)</td>
<td>Retraining programs in Russia and Romania</td>
<td>Registered unemployed adults with income less than 50% of minimum wage, employed 6 of last 12 months or recent graduate; 45% with university degree</td>
<td>Training and unemployment benefits, including a public service component, where local government and other eligible organizations propose public projects and hire ALMP participants to work on them</td>
<td>Statistically significant effects on the likelihood of employment, the likelihood of being employed at least once, and wage levels. -Middle-aged had largest impacts. -Retraining increases the probability of employment and decreases the wage for females. -Program not beneficial for highly educated.</td>
</tr>
<tr>
<td>Bidani, Goh, and O'Leary (2002)</td>
<td>Job training in urban China</td>
<td>Adults</td>
<td>Retraining and job search assistance administered by local labor bureaus to promote labor market entry of people laid off by state-owned enterprises</td>
<td>Positive impact on employment probability in Wuhan. -No effect on earning in Shenyang. -Negative effect on employment probability in Shenyang.</td>
</tr>
<tr>
<td>Cho et al. (2013)</td>
<td>Constraints on women for vocational training in urban Malawi</td>
<td>759 poor vulnerable youth (orphans or school dropouts) aged 15 to 24</td>
<td>Three-month vocational and entrepreneurship program</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Galasso and Ravallion (2004)</td>
<td>Social protection in urban Argentina</td>
<td>Heads of households with dependents who became unemployed as a result of Argentina’s economic crisis in 2003</td>
<td>Participants worked 20 hours per week at a private company and in exchange received direct income support</td>
<td></td>
</tr>
</tbody>
</table>

- Training outcomes were better for male trainees.
- Women faced greater obstacles in undertaking and taking advantage of the training.
- One-third of participants dropped out. Whereas males dropped out to take advantage of unrelated job opportunities, females dropped out more due to external constraints.
- Training more costly to women, who had less access to financing and used more personal savings and were treated worse than men during apprenticeship.
- 26% of participants would have been unemployed and 23% would have been inactive without program.
- Study found substantial leakage to ineligibles, but the program was still well targeted at the poor.

Continued
<table>
<thead>
<tr>
<th>Study</th>
<th>Features</th>
<th>Sample</th>
<th>Intervention</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Galasso, Ravallion, and Salvia (2004) | From workfare to work in urban Argentina      | Adult beneficiaries of temporary employment programs | Skill training and/or vouchers for workfare participants to give to prospective employers (18 month wage subsidy) | - Voucher reduced probability of unemployment (despite fact that few firms made use of the voucher).  
- Private sector employment was 15% for voucher recipients compared to 9% for controls.  
- Women and younger workers had largest impacts. |
| Groh et al. (2012)           | Training and wage subsidies for female youth in Jordan | Young women college graduates               | Awarded voucher to pay employer subsidy equivalent to minimum wage for up to 6 months, if graduate is hired; invited to attend 45 hours of employability skills training designed to provide them with soft skills | - Job voucher led to a 40 percentage point increase in employment in short run, but most employment was not formal.  
- Average effect was smaller and no longer statistically significant 14 months after the voucher period ended.  
- Soft skills component led to sustained improvement in life outlook and reduced depression. |
<table>
<thead>
<tr>
<th>Study</th>
<th>Description</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ibarraran et al. (2012)</td>
<td>Life skills, employability and training for youth in the Dominican Republic</td>
<td>Men and women 16–29 years old from lower socioeconomic strata, and those meeting educational and skill requirements of participating firms</td>
</tr>
<tr>
<td></td>
<td>Vocational classroom and on-the-job training: soft skills</td>
<td>-7%–12% increase in employment for women only.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3–7 hours per week increase in weekly hours worked for women only.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Especially successful in countries with established private vocational training industry, significant wage employment, and high female mobility.</td>
</tr>
<tr>
<td>Jalan and Ravallion (2003)</td>
<td>Anti-poverty program in urban Argentina</td>
<td>Workers affiliated with a successful project and not receiving unemployment benefits from another program</td>
</tr>
<tr>
<td></td>
<td>Up to six month ‘socially useful projects’ submitted by local governments and non-governmental organizations</td>
<td>-Average gain in household income for participants was $103 dollars, with greater gains for younger people.</td>
</tr>
<tr>
<td>Kluve, Lehmann, and Schmidt (1998)</td>
<td>Active labor market policies in Poland</td>
<td>Unemployed workers offered program at their local labor office</td>
</tr>
<tr>
<td></td>
<td>Three forms of training: publicly financed training and retraining, wage subsidies for workers in private or public firms, and public works; 2–3 month courses with unemployment benefits</td>
<td>-Training/program increases average employment probability for women and men over both short and medium term.</td>
</tr>
<tr>
<td></td>
<td>Training/program increases average employment probability for women and men over both short and medium term.</td>
<td>-Non-training ALMP did not have a positive benefit.</td>
</tr>
<tr>
<td>Lubyova and Ours (1999)</td>
<td>Active labor market programs in the Slovak Republic</td>
<td>Adult unemployed workers</td>
</tr>
<tr>
<td></td>
<td>Two ALMP programs providing retraining and counseling services, and wage subsidies at either socially purposeful jobs (up to 2 years) or publicly useful jobs (up to 6 months at public works-type job)</td>
<td>-Workers who enter ALMP have a 150% increase in the exit rate into a regular job.</td>
</tr>
<tr>
<td></td>
<td>Benefits of retraining were only observed for socially purposeful jobs.</td>
<td>-Benefits of retraining were only observed for socially purposeful jobs.</td>
</tr>
<tr>
<td>Study</td>
<td>Features</td>
<td>Sample</td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------------------</td>
<td>-----------------------------------------</td>
</tr>
</tbody>
</table>
| Maitra and Mani (2014)| Learning and earning in urban India | Young women aged 18–39, average age 22 | Six-month vocational training program in stitching and training for young women | -Employment increased by six percentage points.  
-Self-employment increased by four percentage points.  
-Earnings increased by 150% monthly.  
-Positive effects sustained 18 months after completion of training and were attributed to an increase in skills rather than in self-confidence or having a training certificate.  
-More than 40% of trainees did not complete six-month course due to distance to training center and lack of available child care support.  
-Cost-benefit analysis suggests that program is highly cost effective.  
-Using a second round of follow-up data 18 months after the program, all effects are sustained in medium term. |
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Description</th>
<th>Target Group</th>
<th>Training Details</th>
<th>Impacts</th>
</tr>
</thead>
</table>
| Nopo, Robles, and Saavedra (2007) | Gender and racial discrimination in hiring in urban Peru | Technicians and professionals from middle and lower classes in Lima; 80% with above high-school education | Classroom training and internships lasting three months focused on training females for traditionally male occupations; stipend, doubled for child care | - Positive employment impacts for women of 6% at 12 months and 15% at 18 months.  
- Negative impacts for men.  
- After 18 months, beneficiary females generate 93% more labor income than their control counterparts.  
- Decrease in measures of occupational segregation.  
- Program trainees found jobs more quickly.  
- Impacts mainly for trainees older than 25 with work experience.  
- Cost effective for women over 25 but not for younger women. |
| Revenga, Riboud, and Tan (1994) | Retraining in urban Mexico | Youth and adults. Offered to 250,000 registered unemployed people age 20–55, selected on basis of eligibility index. | Short-term vocational education training | - Program trainees found jobs more quickly.  
- Impacts mainly for trainees older than 25 with work experience.  
- Cost effective for women over 25 but not for younger women. |
Note

Mayra Buvinic (corresponding author) is Senior Fellow at the United Nations Foundation; email address: mayra.buvinic@gmail.com. Rebecca Furst-Nichols is Program Officer, United Nations Foundation; email address: rfurst.nichols@gmail.com. This work was supported by a generous grant from the ExxonMobil Foundation to the United Nations Foundation. The authors are indebted to Noa Gimmelli, ExxonMobil Foundation, and Emily Courey Pryor, United Nations Foundation, for their steadfast support.


References


Employer Voices, Employer Demands, and Implications for Public Skills Development Policy Connecting the Labor and Education Sectors

Wendy V. Cunningham and Paula Villaseñor

Educators believe that they are adequately preparing youth for the labor market while at the same time employers lament the students’ lack of skills. A possible source of the mismatch in perceptions is that employers and educators have different understandings of the types of skills valued in the labor market. Using economics and psychology literature to define four skills sets—socio-emotional, higher-order cognitive, basic cognitive, and technical—this paper reviews the literature that quantitatively measures employer skill demand, as reported in a preference survey. A sample of 27 studies reveals remarkable consistency across the world in the skills demanded by employers. While employers value all skill sets, there is a greater demand for socio-emotional skills and higher-order cognitive skills than for basic cognitive or technical skills. These results are robust across region, industry, occupation, and education level. Employers perceive that the greatest skills gaps are in socio-emotional and higher-order cognitive skills. These findings suggest the need to re-conceptualize the public sector’s role in preparing children for a future labor market. Namely, technical training is not equivalent to job training; instead, a broad range of skills, many of which are best taught long before labor market entry, should be included in school curricula from the earliest ages. The skills most demanded by employers—higher-order cognitive skills and socio-emotional skills—are largely learned or refined in adolescence, arguing for a general education well into secondary school until these skills are formed. Finally, the public sector can provide programming and incentives to non-school actors, namely parents and employers, to encourage them to invest in the skills development process. JEL codes: J23, J24
Recent years have seen a rethinking of the nature of, and measurement of, skills related to labor market success, which increasingly draws on concepts developed in the psychology literature. Traditionally, the economics literature has focused on “years of schooling” as a proxy for these skills, but Hanushek and Woessman (2008) argue that this approach suffers from serious measurement error and should be discarded in favor of exploring the role that particular skills play in driving labor market outcomes.1 James Heckman and others parse the concept of “skills” into “cognitive” skills—such as knowledge, comprehension, and critical thinking—and “non-cognitive” skills, roughly defined as personality traits and socio-emotional behaviors.2 This distinction is increasingly relevant in the context of modern labor markets where jobs increasingly involve non-routine analytical tasks that require problem-solving, emotional intelligence, logic, and teamwork, rather than traditional routine manual tasks (Acemeglu and Autor 2011; Autor 2013; Aedo et al. 2013; World Bank 2015).3

At the same time, emerging evidence suggests that schools are not teaching the skills that labor markets need. A 37-country study reports that about one-third of the 35,000 employers interviewed have trouble finding workers with the right skills (Manpower 2010). Moursched et al. (2012) interviewed nearly 3,000 employers and 1,000 education providers in nine countries, and reports that 42% of employers believe that graduates are prepared for the labor market as compared to 72% of educators. While the recurring shortfall in cognitive skills as measured, for instance, by PISA scores, has been convincingly documented, an emerging body of literature suggests that we must cast the skills net wider. Shivpuri and Kim (2004), for instance, find that U.S. employers report that they most value interpersonal skills, followed by trustworthiness, leadership, and perseverance, while educators identify “knowledge” as the most important skill.

In light of these emerging views, this paper systematically reviews the evidence on the skills demand profile of employers in developed and developing countries, with the objective of informing education and training policy.4 Specifically, the paper reviews 27 studies that explicitly ask employers to provide their ranking of the most important skills for their workforce, and to provide a ranking of the skills that are most lacking. The paper then utilizes a developmental psychology framework to recommend shifts to public policies and programs to bring the supply and demand for skills closer together.

Skills for the Labor Market: Concepts and Definitions

Following the literature on economics, as well as on behavioral, personality, and industrial and organizational (IO) psychology, we classify labor-market related skills into two groups: cognitive and socio-emotional.5 The American Psychological
Association (APA) defines cognitive skills as the “ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles [through deliberate thinking],” (Neisser et al. 1996). This definition includes intelligence, reasoning, information-processing, perception, memory, literacy, numeracy, and learning, among others. Within cognitive skills, we define three sub-sets for the purposes of this paper: basic cognitive, higher-order cognitive, and technical.

Basic cognitive skills encompass fundamental academic knowledge and comprehension, including literacy and math, and are strongly correlated with labor market success. Studies that use longitudinal data from the US to regress labor force variables of young adults on cognitive test scores collected while the sample was in high school, find that a one standard deviation increase in a mathematics test score in 12th grade is correlated with 10–15% higher annual earnings by the mid-twenties to early thirties (Mulligan 1999; Murnane et al. 2000; Lazear 2003). Similar results are found for the UK (McIntosh and Vignoles 2001) and Canada (Finnie and Meng 2001). Using literacy scores and labor force behaviors from the International Adult Literacy Survey (IALS), Hanushek and Zhang (2009) finds that a one standard deviation increase in literacy scores increases earnings by 9.3% in a 13 country sample. A small set of papers find an impact of cognitive skills on wages in developing countries as diverse as Ghana (Glewwe 1996), Kenya (Knight and Sabot 1990), Pakistan (Alderman, Behrman, Ross and Sabot 1996), and South Africa (Moll 1998). These skills are the core of school curricula and are applicable across a wide range of professions.

Higher-order cognitive skills encompass the capacity to deal with complex information processing and are aligned with the emerging non-routine job tasks noted above. These tasks include such abilities as critical thinking, application of knowledge, analysis, problem solving, evaluation, oral and written communication, and adaptive learning (Bloom et al. 1956). An emerging body of literature shows that these skills drive labor market outcomes across a wide-range of occupations (Herrnstein and Murray 1994; Hanushek and Woessmann 2008). For example, Gottfredson (1997) argues that reasoning, problem solving, and decision-making are among the most powerful predictors of overall job performance. These skills are often not included in traditional school curricula (Pina et al. 2012).

Technical skills can be defined as the specific knowledge required to carry out an occupation, and are often equated with “job skills” in policy circles. Such skills might include the ability to repair a car’s muffler, being able to identify specific bacteria under a microscope, or the know-how to sew dozens of shirts per hour. While the definition of technical skills overlaps with that of basic- and higher-order cognitive skills, they merit separate treatment in a policy-oriented paper in order to test the assumption that they alone are synonymous with job-skills, and to provide insights for refining job training policy. A review by Betcherman et al. (2007) of
job training programs across the world finds, at best, positive returns to technical training for women and, in some cases, youth. More commonly, technical training programs yield zero, or negative rates of return, the latter indicating that more valuable skills would have been acquired if that person had spent her time working rather than in a training course. These skills may be included in the school curriculum, or may be taught through vocational and technical training programs or institutes.

Socio-emotional skills, the second broad group of labor market-related skills, are behaviors, attitudes, and traits that are necessary complements to cognitive skills in the production process. These skills include such diverse concepts as grit to finish a job, working in teams, commitment, creativity, and honesty. A review of meta-evaluations, reported in Almlund et al. (2011), confirms a significant correlation between job performance and openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. Individual studies find conditional correlations between socio-emotional skills and wages (Heckman and Rubenstein 2001; Kern et al. 2013; Mueller and Plug 2006; Heckman, Stixrud, and Urzúa 2006), professional success (Duckworth et al. 2007), absenteeism in the workplace (Stormer and Fahr 2010), and a range of other labor force outcomes. In fact, these skills may be more important than cognitive skills for employment outcomes, especially among low-skilled occupations (Bowles, Gintis, and Osborne 2001; Mueller and Plug 2006; Carneiro, Crawford, and Goodman 2007; Kniesner and ter Weel 2008). These skills are not commonly a part of the standard school curriculum (Pina et al. 2012).

While this evidence may convincingly signal that a range of cognitive and socio-emotional skills are important for labor market success, it is not likely to capture the employer perspective for a few reasons. First, supply-side surveys measure only a small set of pre-defined skills, thereby potentially missing the skills that employers most value. Second, the supply-side data measure the skills that the current labor force has, thereby not capturing skills that may not be available, but may be important for the production process. This points to a need to complement the existing literature with demand-side data measuring the skills that employers want and the extent to which these skills demands are met by the current education and training systems.

Data

This paper draws from the published literature across a range of fields. Each data point represents the statistical finding of a paper that summarizes raw data from employer surveys, most of which were specialized surveys collected for the purpose of understanding employer skill needs.

Using a large set of key words related to employer demand for skills, we search five databases to build our sample: Google Scholar, IDEAS-RePEc, EconLit, ERIC (Education Resources Information Center), and PsyContent. In addition,
we consult the publications of institutions dedicated to research in the education field such as the National Center on Education and the Economy (NCEE) and the International Association for the Evaluation of Educational Achievement (IAE). We follow a snowball approach by searching the citations of relevant papers we found in the demand-for-skills literature, using a similar approach to improve the key words list.

We select the studies to be included in the sample based on four criteria. First, we limit the sample to studies published from 2000–2015 in order to capture those skills that are likely relevant for today’s policymakers while providing a sufficiently large sample of studies from various countries to carry out the meta-analysis. Second, we only include studies that present results that summarize empirical data collected from employers, thereby excluding studies based on expert opinion or qualitative assessments. We make one exception by including an analysis of the US O*NET (Casner-Lotto and Barrington 2006), which tabulates data provided by workers. Third, we limit the sample to studies that ask employers to rank the demand for, or difficulty in acquiring all four of, our broad skills sets—basic, higher-order, technical, and socio-emotional—in order to provide a preference ranking across the skill sets. Studies that do not provide a ranking or do not disaggregate within these broad skills sets are included in the discussion but not in the statistical meta-evaluation (Bassi et al. 2012; Burnett and Jayaram 2012; IFC 2011). Fourth, we attempt to ensure a reasonable level of comparability in the structure of the questions across the sample.

Our approach yields a sample of 27 studies drawn from data from dozens of countries. The sample includes two global studies (9 countries in Moursed et al. 2012, 40 countries in Andreasson 2009), three studies from Western Europe (all in the UK), five from Eastern Europe (Romania, Russia, Macedonia, Poland, and Kazakhstan), five from Latin America and the Caribbean (St. Kitts, Peru, Mexico, and two that survey employers in several LAC countries), six from South/East Asia and the Pacific (India, Pakistan, Indonesia, Vietnam, Tonga, Solomon Islands), four studies from the United States, and one each from Africa (Botswana) and the Middle East (Lebanon). About half the studies select a nationally representative sample, stratifying by firm size, industry, or other characteristics, while others use ad hoc sampling methods such as drawing a sample from another survey (World Bank 2012; Shivpuri and Kim 2004; and Beneitone et al. 2007) or over-sampling specific types of workers (Bassi et al. 2012 and CIDAC 2014 focus on youth-oriented industries; Burrus et al. 2012 selects higher-skilled occupations). The sample principally draws from small, medium, and large firms operating in the formal sector, while a few include a sizeable number of micro-firms (Martin et al. 2008; Hamid et al. 2011; Rutkowski 2010). A wide range of industries are represented, though the manufacturing sector is over-represented. While most of the studies ask employers about their entire labor force, some ask employers to focus on certain groups, such
as youth (Mourshed et al. 2012; CBI 2012; Martin et al. 2007; Close 2012; IFC 2011; Shivpuri and Kim 2004) or new recruits (Casner-Lotto and Barrington 2013; Zemsky 1997).

Due to our selection criteria, the questions of interest are similarly structured across the sample. Twenty four studies ask employers to identify the most important skills that they need in their employees, either by asking an open question or by providing a list of pre-determined skills, and to rank the skills they listed in order of importance. We refer to the results from these studies as the “priority ranking” of the most important skills. Fifteen studies measure the most important skills gaps. Employers are asked which skills are most lacking and to rank them according to how large the perceived gap is. We refer to the results of these studies as the “gaps ranking”. Each study present rankings for all employers in its sample, and a few also disaggregate rankings by industry, occupation, or worker education level.

The broad range of countries potentially represent different labor market characteristics and institutions. This could complicate comparisons if different results emerge in different kinds of labor markets. However, the findings of the studies in our sample are quite similar, suggesting robust patterns that cut across labor market structures.

The skills reported in the studies in our sample are stated preferences rather than revealed preferences, but several studies suggest that there is a strong overlap between the two. Globally, employers are increasingly using headhunter firms and IO psychologists to design and administer skills tests for hiring and promotion (Salgado et al. 2001); the tests typically include the four broad skills sets that we have defined for this paper. Since employers pay IO psychologists to administer these tests, we may assume that the measured skills are valued by the employer. A comparison of a survey to assess employer skill demand and the hiring tests used by headhunters finds a significant overlap between the two surveys (CIDAC 2014). Coming at the question from a different angle, the US O*Net, which asks employees to assess the degree to which they use different skills in nearly 1,000 professions, reports the use of skills classified in all four of our broad skill sets (Burrus et al. 2013), thereby demonstrating that skills actually used on the job are quite similar to those measured in the employer demand surveys used for this paper.

The skills that employers value, as reported in the studies in our sample, are closely matched to those that they attempt to detect when hiring, whether for low- or high-skilled workers. Referring back to the IO psychology literature, hiring tests are regularly administered to potential new low- and high-skilled workers (Salgado et al. 2001). Further, a review of the skills posted on a Ukrainian public job-search site reports that socio-emotional skills are required of low- and high-skilled jobs, while technical skills are only specified for more skilled jobs (World Bank forthcoming). This is consistent with Osterman (2001), who concludes that employers of low-skilled workers in the US particularly depend on socio-emotional skills, which
are congruent with low-paid service jobs, given the very low technical demands of the job, while employers of higher-skilled workers put greater emphasis on establishing technical skills, although employers seek IO psychology testing in all three skill sets (Anderson 2001). While socio-emotional and higher-order cognitive skills may be difficult to assess, particularly in low-skilled workers, this may be where first impressions in the interview and employer preference for those with work experience play a role.20

Employer responses are qualitative and not quantitative, so while we know which skill is most important and which skills gap is more pressing, we do not know how much more important a skill is as compared to a skill ranked just below it (Rutkowski 2010). To address this, we not only present the skill that is rated as the most important for each study, we also consider the second-most important skill and the five most important skills (those ranked first through fifth). We report the first, second, and top five ranked skills gaps as well.

When employers identify the most important skill, they may not consider skills that are crucial to the production process, but are abundant in the work force. Since these skills are not used as a selection mechanism when hiring, assigning wages, or awarding promotions, they may be inadvertently forgotten by those being surveyed. To ensure that employers are not only reporting skills gaps, several of the surveys differentiate between the most important skill and the most important skill gap. Results presented below find that the results are, indeed, different.

Methodology

To identify those skills most demanded by employers and the biggest skills gaps, we first map the skills identified in each study into one of our four broad skills sets, \( s = \{ \text{basic cognitive, higher-order cognitive, technical, and socio-emotional} \} \), as defined in table 1.21,22 We then calculate the share of studies \( n = 1, \ldots, N \) that identify \( s \), for rank \( r \)

\[
\frac{\sum_{n=1}^{N} s_{n} = s|_{r,s}}{N}.
\]

We start with \( r = 1 \) and then repeat the exercise for the second, third, fourth, and fifth priority rankings \( (r = 1, \ldots, 5) \), giving us a matrix of skill set-ranking cells.23 Finally, we calculate the frequency that each skill set is listed in the top five. Since employers can select any skill for any \( r \), the same broad skill set \( s \) may appear several times in the top five ranking. We repeat the exercise with employer rankings of the biggest skills gap.
<table>
<thead>
<tr>
<th>Socio-emotional</th>
<th>Higher-order cognitive</th>
<th>Basic cognitive</th>
<th>Technical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>Analysis skills</td>
<td>Basic literacy</td>
<td>Advanced IT</td>
</tr>
<tr>
<td>Collaboration</td>
<td>Critical thinking</td>
<td>Numeracy</td>
<td>Advanced vocational</td>
</tr>
<tr>
<td>Commitment</td>
<td>Decision-making</td>
<td></td>
<td>Basic vocational</td>
</tr>
<tr>
<td>Control emotions</td>
<td>Entrepreneurship</td>
<td></td>
<td>Business awareness</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>Foreign language</td>
<td></td>
<td>Capable</td>
</tr>
<tr>
<td>Cooperation</td>
<td>Innovation</td>
<td></td>
<td>Computer literacy</td>
</tr>
<tr>
<td>Creativity</td>
<td>Intellect</td>
<td></td>
<td>Degree level</td>
</tr>
<tr>
<td>Conflict aversion</td>
<td>Language</td>
<td></td>
<td>Degree subject</td>
</tr>
<tr>
<td>Cultural diversity</td>
<td>Learning processes</td>
<td></td>
<td>Education level</td>
</tr>
<tr>
<td>Customer awareness</td>
<td>Listening skills</td>
<td></td>
<td>Experience</td>
</tr>
<tr>
<td>Customer care</td>
<td>Manage risk</td>
<td></td>
<td>Grades</td>
</tr>
<tr>
<td>Dependability</td>
<td>Oral communication</td>
<td></td>
<td>Hands-on training</td>
</tr>
<tr>
<td>Detail-oriented</td>
<td>Organization</td>
<td></td>
<td>Industry-based skills</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Planning</td>
<td></td>
<td>IT knowledge</td>
</tr>
<tr>
<td>Emotional stability</td>
<td>Problem-solving</td>
<td></td>
<td>Job-specific skills</td>
</tr>
<tr>
<td>Enterprising</td>
<td>Strategic management</td>
<td></td>
<td>Office administration</td>
</tr>
<tr>
<td>Extraversion</td>
<td>Time management</td>
<td></td>
<td>Practical knowledge</td>
</tr>
<tr>
<td>Flexibility</td>
<td>Thinking skills</td>
<td></td>
<td>Professional skills</td>
</tr>
<tr>
<td>Hard worker</td>
<td>Written-communications</td>
<td></td>
<td>Score on employer test</td>
</tr>
<tr>
<td>Honesty</td>
<td></td>
<td></td>
<td>Statistical analysis</td>
</tr>
<tr>
<td>Initiative</td>
<td></td>
<td></td>
<td>STEM</td>
</tr>
<tr>
<td>Independence</td>
<td></td>
<td></td>
<td>Technical skills</td>
</tr>
<tr>
<td>Integrity</td>
<td></td>
<td></td>
<td>Theoretical training</td>
</tr>
<tr>
<td>Leadership</td>
<td></td>
<td></td>
<td>University attended</td>
</tr>
<tr>
<td>Modesty</td>
<td></td>
<td></td>
<td>Work experience</td>
</tr>
<tr>
<td>Motivation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negotiating</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negotiate conflict</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Networking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open to new ideas</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perseverance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal appearance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive attitude</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proactive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Punctuality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professionalism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responsibility</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-confidence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-management</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social values</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stress-management</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teamwork</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work ethic</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Cunningham and Villaseñor
Results are presented in three ways. First, the broad skills sets corresponding to the five most important skills listed in each of 24 studies are discussed. Second, the skill set that was ranked as most important is presented, as well as a discussion of the skills underlying those broader sets. Finally, the second-most important skill set is briefly noted. We also compare the priority skills demanded by employers in different regions, industries, occupations, and by worker education level. We repeat this exercise using data from the 15 studies that measure the greatest skills gaps: presentation of the five greatest skills gaps, the most important skills gap, the second-most important skill gap, and a comparative discussion by region and occupation.24

Results

Skills Employers Most Value

The top panel of table 2 shows that of the five skills ranked as most important, more than half can be classified as socio-emotional, another 29.7% as higher-order cognitive, and 12.7% as technical. Specifically, socio-emotional skills are named 60 times from a potential 118 responses. Higher-order cognitive skills are listed 35 times, and technical skills 15 times. Basic cognitive skills are only named eight times among the top five preferred skills in our 24-study sample.

Limiting the responses to only the skill identified as the most important, 79% name socio-emotional skills (table 2, top panel, ranking “1”). Nineteen studies rank

| Table 2. Employer Ranking of Most Important Skills and Greatest Skills Gaps, Percentage |
|----------------------------------|-------------------|-----------------|---------------|-------------|-------------|
| Ranking       | Socio-emotional | Higher-order cognitive | Basic cognitive | Technical | Sample size |
| Most important skill |                  |                   |               |            |             |
| 1             | 79.2             | 16.7              | 0.0           | 4.2        | 24          |
| 2             | 41.7             | 37.5              | 12.5          | 8.3        | 24          |
| 3             | 45.8             | 29.2              | 8.3           | 16.7       | 24          |
| 4             | 45.8             | 33.3              | 4.2           | 16.7       | 24          |
| 5             | 40.9             | 31.8              | 9.1           | 18.2       | 22          |
| Total         | 50.8             | 29.7              | 6.8           | 12.7       | 118         |

Greatest skill gap

<table>
<thead>
<tr>
<th>Ranking</th>
<th>Socio-emotional</th>
<th>Higher-order cognitive</th>
<th>Basic cognitive</th>
<th>Technical</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.0</td>
<td>46.7</td>
<td>0.0</td>
<td>33.3</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>53.3</td>
<td>26.7</td>
<td>0.0</td>
<td>20.0</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>53.3</td>
<td>33.3</td>
<td>6.7</td>
<td>6.7</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>60.0</td>
<td>26.7</td>
<td>0.0</td>
<td>13.3</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>21.4</td>
<td>57.1</td>
<td>7.1</td>
<td>14.3</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>41.9</td>
<td>37.8</td>
<td>2.7</td>
<td>17.6</td>
<td>74</td>
</tr>
</tbody>
</table>

Source: Authors' elaboration based on sample data.
a socio-emotional skill as their first priority, with particular emphasis on teamwork, honesty, and punctuality, but also listing work ethic, interpersonal skills, work attitude, integrity, life skills (negotiation, cultural diversity), and responsibility (table 3). Another four studies rank higher-order cognitive skills as the most important skill set, including communications, problem-solving, and critical thinking skills. One study identifies technical skills as the most important skill set, though the specified “job-related skills” may encompass a larger set of skills. No study names basic cognitive skills as the most important skill set.

Considering the skill identified as the second most-important, socio-emotional skills again emerge the most frequently (table 2, top panel). More than 40% of the sample names a socio-emotional skill, listing many of the same skills that are ranked first. Another 37.5% lists a higher-order cognitive skill, adding time management to the list of priority skills. Basic cognitive skills are named by 12.5% of the sample, with an emphasis on literacy, while 8% identify a technical skill as the second priority. Again, the technical skills named—practical technical skills and professional knowledge—can also encompass other broad skills sets.

These results are consistent with two studies that survey employer preferences in several countries; namely that socio-emotional skills are most important. Moursheid, Farell, and Barton (2012) interview employers in nine countries and ask them to rate, on a scale from one to ten (low to high) the importance of 13 predetermined skills; 80% rank work ethic or teamwork—both socio-emotional skills—as the top skills. The higher-order cognitive skill of language and oral communication is ranked third (72% of employers), and hands-on training in discipline (technical) is ranked fourth, with approximately 70% of employers citing it as very important. Andreasson (2009) finds similar results among business executives in 40 countries who identify that the skills most in demand in the next decade are “life skills”, defined as negotiating, networking, working with cultural diversity (48% of the sample), followed by problem solving and leadership.

The equation (1) results for the 24-study sample may mask heterogeneity among the skills needs of employers in different contexts. Thus, we cut the data along different dimensions to explore if the aggregate results are reflected in sub-sets of the sample.

**Skills Preferences by Region**

Developed countries, with their more sophisticated jobs, may have a greater need for higher-order cognitive and socio-emotional skills compared to less-complex developing country economies. However, different regions of the world show similar patterns in the broad skill sets that employers most value. Socio-emotional skills most frequently emerge among the five most important skills in each region. The only exception is the Middle East, where higher-order cognitive skills dominate in
<table>
<thead>
<tr>
<th>Region/country</th>
<th>Priority ranking</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>World</td>
<td>work ethic</td>
<td>Moursheed et al. 2012</td>
</tr>
<tr>
<td>World</td>
<td>life skills</td>
<td>Andreasson 2009</td>
</tr>
<tr>
<td>LAC</td>
<td>interpersonal</td>
<td>Beneitne et al. 2007</td>
</tr>
<tr>
<td>LAC</td>
<td>critical thinking</td>
<td>Ogier 2009</td>
</tr>
<tr>
<td>Botswana</td>
<td>hard work</td>
<td>World Bank 2014</td>
</tr>
<tr>
<td>India</td>
<td>integrity</td>
<td>Blom and Saeki 2011</td>
</tr>
<tr>
<td>Indonesia</td>
<td>thinking skills</td>
<td>diGropello 2011</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>independent</td>
<td>Sondergaard 2012</td>
</tr>
<tr>
<td>Lebanon</td>
<td>teamwork</td>
<td>World Bank 2012</td>
</tr>
<tr>
<td>Macedonia</td>
<td>responsibility</td>
<td>Rutkowski 2010</td>
</tr>
<tr>
<td>Mexico</td>
<td>teamwork</td>
<td>CIDAC 2014</td>
</tr>
<tr>
<td>Pakistan</td>
<td>punctuality</td>
<td>Hamid et al. 2011</td>
</tr>
<tr>
<td>Peru</td>
<td>teamwork</td>
<td>World Bank 2011</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>World work ethic</td>
<td>teamwork</td>
</tr>
<tr>
<td>local language</td>
<td>communication</td>
</tr>
<tr>
<td>hands-on experience</td>
<td>teamwork</td>
</tr>
<tr>
<td>critical thinking</td>
<td>leadership</td>
</tr>
<tr>
<td>technical</td>
<td>learning processes</td>
</tr>
<tr>
<td>life skills</td>
<td>Leadership</td>
</tr>
<tr>
<td>Adaptable</td>
<td>communication</td>
</tr>
<tr>
<td>willingness to learn</td>
<td>entrepreneurship</td>
</tr>
<tr>
<td>Computer</td>
<td>n/r</td>
</tr>
<tr>
<td>literacy</td>
<td>analytical skills</td>
</tr>
<tr>
<td>team work</td>
<td>time management</td>
</tr>
<tr>
<td>communication</td>
<td>problem solving</td>
</tr>
<tr>
<td>customer care</td>
<td>motivation</td>
</tr>
<tr>
<td>innovation</td>
<td>conflict resolution</td>
</tr>
<tr>
<td>reliability</td>
<td>behavior</td>
</tr>
<tr>
<td>proactive</td>
<td>work under pressure</td>
</tr>
<tr>
<td>Country</td>
<td>Skill 1</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Poland</td>
<td>responsibility</td>
</tr>
<tr>
<td>Romania</td>
<td>efficiency</td>
</tr>
<tr>
<td>St Kitts</td>
<td>honesty</td>
</tr>
<tr>
<td>Solomon Islands</td>
<td>work attitude</td>
</tr>
<tr>
<td>Tonga</td>
<td>honesty</td>
</tr>
<tr>
<td>Vietnam</td>
<td>job-related skills</td>
</tr>
<tr>
<td>UK</td>
<td>communication</td>
</tr>
<tr>
<td>US</td>
<td>interpersonal</td>
</tr>
<tr>
<td>US</td>
<td>problem solving</td>
</tr>
<tr>
<td>US</td>
<td>work ethic</td>
</tr>
<tr>
<td>US</td>
<td>attitude</td>
</tr>
</tbody>
</table>

Source: Data drawn from each paper cited in last column; “n/r” indicates that a ranking was not requested.
the single-country study from that region (World Bank 2012). Higher-order cognitive skills appear second-most frequently among the top five skills in Latin America and the Caribbean (LAC), Eastern Europe and developed countries, while technical skills are the second-most cited in the East Asia sample. Basic cognitive skills are rarely mentioned among the top five skills in any region.

Socio-emotional skills are the top-rated skill set in most studies in LAC, South East Asia, Middle East/North Africa, Africa, Eastern Europe, and developed countries (US and UK), and the specific type of skill is similar across studies. Five of the six Latin American surveys conclude that employers most value socio-emotional skills, specifically, inter-personal skills, teamwork, attitude, responsibility, honesty, and integrity (Beneitone et al. 2007, Bassi et al. 2012, World Bank 2011, Blom and Hobbs 2008). The two South East Asia studies give top ranking to socio-emotional skills, that is, punctuality, honesty, integrity, and reliability (Hamid et al. 2011, Blom and Saeki 2011). The sole Middle East/North Africa study reports that 70% of Lebanese employers most highly rate teamwork (World Bank 2012) while work ethic is the top rated skill by 70% of Botswanan employers, reported in the only African study (World Bank 2014). Three of the four Eastern European studies rank a socio-emotional skills as the priority skill, singling out responsibility and efficiency as the most desired skills in Romania, Macedonia, and Poland (Balcar 2012, Rutkowski 2010, Arnhold et al. 2011); a higher-order cognitive skill (independent work) is ranked first by nearly all employers in Kazakhstan (Søndergaard et al. 2012). Three of the five developed country studies (Shivpuri and Kim 2004; Casner-Lotto and Barrington 2006; Zemsky 1997) find that socio-emotional skills—work ethic, attitude and interpersonal skills—are the preferred skills; the other two countries name higher-order cognitive skills (Burrus et al. 2013, Martin 2008).

The pattern is less uniform in East Asia and the Pacific. Two studies rank socio-emotional skills first, specifying work attitude, and honesty and punctuality (Close 2012; Tonga Chamber of Commerce 2010); this is similar to the skills named in the other regions. However, “job-related skills,” which we classify as technical, were prioritized in the Vietnam study (World Bank 2008), while cognitive skills (higher-order for managers, basic for workers) were ranked first in Indonesia (di Gropello 2011).

Skills preferences are less uniform across regions when assessing the second-most important skill. While socio-emotional skills still dominate among LAC and South Asian employers, other skill sets emerge in the other regions. For example, of the four East Asia Pacific studies, each identifies a different skill set as second priority: basic cognitive in Indonesia (di Gropello 2011), technical in Vietnam (World Bank 2008), higher-order cognitive in the Solomon Islands (Close 2012), and socio-emotional in Tonga (Tonga Chamber of Commerce 2010). Similarly, in Eastern Europe technical skill emerges as a second priority in Romania (Balcar 2012) and basic cognitive
(literacy) is identified in Macedonia (Rutkowski 2010), while socio-emotional skills (motivation and time management) are ranked second in Poland (Arnhold et al. 2011) and Kazakhstan (Sondergaard 2012).

In conclusion, in spite of the small number of studies per region, the evidence to date suggests that all four skills sets are important to employers regardless of region, with more weight placed on socio-emotional and higher-order cognitive skills.

**Most Important Skill by Industry: Manufacturing v. Services**

Given the different nature of job tasks in the manufacturing and service industries, we might expect to find a greater demand for technical skills in the former, and for socio-emotional skills in the latter. Three of our studies separately report the skills preference rankings for employers in the manufacturing sector and those in the service sector (World Bank 2008; diGropello 2011; diGropello 2008), allowing us to explore this assumption. Starting with the top five ranked skills, socio-emotional and higher-order cognitive skills are named most frequently in both the manufacturing and service sectors. Socio-emotional skills were listed six times each (of a potential 18 mentions for each country-industry matrix) for the manufacturing and service sectors, with particular emphasis placed on independent work, teamwork, and creativity in both sectors. Higher-order cognitive skills—problem-solving and communications—were listed six times for the manufacturing sector, and five times in the service sector. Technical skills were named three times each by both sectors: computer skills, practical experience, and theoretical knowledge of the job.

Turning to a comparison of the skill that is ranked $r$ for each industry, the preferred skill set in manufacturing is also the preferred skill set in services in each of the three countries in the sample. For example, Filipino manufacturing employers rank first the socio-emotional skill of independent work (15%). Service sector employers also rank independent work as the most important skill. The pattern is repeated for the second, third and fourth rankings in the countries in the sample.

Although the industry-specific ranking of the most important skills within country are identical, the weight on each priority ranking reveals expected patterns: socio-emotional skills are more important in the service sector than the manufacturing sector. For example, 54% of Indonesian manufacturing sector employers rate behavioral skills as very important (for their professional staff), as compared to 72% of service firm employers (diGropello 2011). Higher-order communications skills are slightly more important in the service sector than the manufacturing sector, as observed in the Philippines (11.5% versus 10%) (diGropello 2010) and Vietnam (10% versus 8%) (World Bank 2008).

There is not a global pattern in the most important skill in each industry across the three studies that provide industry-level disaggregation. For example, employers in the manufacturing sector prioritize socio-emotional skills in Latin America.
(Bassi et al. 2012), “practical” technical skills in Vietnam (World Bank 2008), and cognitive skills in Indonesia (diGropello 2011). Similar patterns emerge for service sector employers. We can conclude, with the caveats associated with a small sample, that within-country skills profiles dominate industry skills profiles.

Most Important Skills by Occupation: Managers v. Workers

Given the leadership and team management role of managers, we might expect employers to demand more higher-order cognitive and socio-emotional skills from their managers and more information-based skills (technical and basic cognitive) from their workers. Six studies report employer rankings of the most important skills for managers as compared to workers (Bassi et al. 2012; CIDAC 2014; diGropello 2011; World Bank 2014; Vasiliev et al. 2013; World Bank 2012). Socio-emotional skills strongly emerge among the top five skills that employers seek in both managers and workers, with higher-order cognitive skills also appearing frequently. Socio-emotional skills were listed ten times (of a potential 30 responses) for managers and 15 times for workers. Teamwork, honesty, hard work, punctuality, and leadership were named for both managers and workers, with slightly more emphasis placed on flexibility, independent work, and self-management for workers. Higher-order skills were listed ten times for managers and six times for workers. Employers prioritize analytical skills, problem-solving, decision-making, and communication among managers, with more emphasis placed on communication for workers. Technical and basic cognitive skills were named once each for managers and for workers.

Employers across countries demand similar skills by occupation, though the range of preferred skills of workers is much broader than that of managers. In all six countries in the sample, employers sought managers with strong socio-emotional and higher-order cognitive skills, particularly problem-solving, teamwork, and leadership. These skills were also demanded of workers across countries, with teamwork added to the list in Russia, Lebanon, and Mexico (Vasiliev et al. 2013; World Bank 2012; CIDAC 2014). However, a range of additional skills were sought in workers, such as basic cognitive skills (diGroppelo 2011), technical skills (Vasiliev et al. 2013), communications (World Bank 2012; CIDAC 2014), punctuality (World Bank 2014; World Bank 2012), and hard work (World Bank 2014). This small sample comparison suggests that there is a global set of skills for managers, while worker skill sets are somewhat country specific.

Most Important Skill for the Less v. More Educated Workforce

Since more educated workers are likely to hold more complex jobs, employers may require more higher-order cognitive skills and technical skills for these employees than for those who are less educated. Four country studies report employer demand
for skills for more educated compared to less educated employees (Burrus et al. 2013; Casner-Lotto and Barrington 2006; World Bank 2011, 2008), and find that employers overwhelmingly prioritize socio-emotional skills across education levels, while valuing higher-order cognitive skills in their more educated workers. Socio-emotional skills were listed 13 times (of a potential 20 responses) as one of the top five skills of less educated workers as compared to three mentions of a technical skill and one each of higher-order and basic cognitive skills. Similarly, socio-emotional skills were listed 10 times for more educated workers, and higher-order skills were listed five times. Technical skills were named twice as one of the top skills among more educated workers.

Within country comparisons reveal that employers give a similar priority ranking to the broad skill sets sought in more and less educated workers, though the specific skill differs by education level in the four countries in the sample. For example, of the seven skills presented in the Peruvian survey, employers most demanded socio-emotional skills for more and less educated workers, naming interpersonal skills for workers with more than a tertiary education (17%) and teamwork for workers with less than a completed secondary education (23%; World Bank 2011).

Cross-country comparisons show that regardless of education level, workers are expected to excel in socio-emotional skills. Employers of less educated workers in the four country sample all prioritize socio-emotional skills, naming work ethic, dependability, teamwork, and punctuality. The same skills are listed by employers of more educated workers in all four countries. The only exceptions are that higher-order cognitive (communication) emerges for more educated workers in one country (the United States, as reported in Casner-Lotto and Barrington 2006), and Vietnam emphasizes technical skills for both its more and less educated workers, while it is barely mentioned by employers in other countries.

**Partial Correlations**

The Indonesia study presents a particularly rich disaggregation of employer skills demands, allowing us to understand priority skills in finer segments of the labor market.26 The data show that the most important skills that employers seek in managers are similar whether the manager is in the service or in the manufacturing sector. The only trend that is dampened when aggregating the data by occupation (such that industry is no longer broken out) is that the share of employers that identify socio-emotional skills as most important in their managers (64%) is overstated for manufacturing managers (54%) and understated for service managers (72%); however, this skill is ranked as second-most important for managers regardless of sector. The most important skill for workers are, again, quite similar across industries with one exception: socio-emotional skills are ranked second by service sector employers (38%) and third by manufacturing sector employers (25%; diGropello 2011).
Greatest Skill Gaps

As noted above, it is possible that when employers cite the most important skill for the production process, they may actually be citing the greatest skill gap. A review of the skills gap literature shows that employers are not confusing these concepts, with different trends emerging in the gap analysis than those observed in the previous exercise.

The bottom panel of table 2 presents the greatest skills gaps, as calculated in equation (1), and shows that higher-order cognitive skills emerge more strongly in the gap analysis than in the priority ranking of the most important skills. More than 40% of the 15-study sample identifies a socio-emotional skill as one of the top five skill gaps, naming a lack of behavioral and inter-personal skills, whereas “work ethic,” which emerged strongly as one of the most important skills valued by employers, as discussed above, does not emerge as a priority skill gap (table 4). Nearly an equal share of studies, 38%, name a higher-order cognitive skill as one of the top five skill gaps. The specific skills listed are similar to those identified above by employers as the most important skills.

Turning to the top-ranked skill gap, higher-order cognitive skills are named by nearly half of the employers, and technical skills, including professional skills, job-specific skills and work experience, are named by 33% of the sampled employers. While we classified all of these as “technical”, several likely include all four of our broad skills sets. For example, “professional skills” or “work experience” may include knowledge of specific equipment (technical), working with others (socio-emotional), the ability to resolve problems (higher-order cognitive), and basic math for operating the equipment (basic cognitive).

The second-most important skills gap is dominated by socio-emotional skills, named by 53.3% of the sample (table 2, bottom panel). Behavioral rather than values-related skills (such as honesty, work ethic) emerge. Another 26% name higher-order cognitive skills as the second-most important skill gap, while one-fifth of the sample lists a technical skill. Basic cognitive skills are barely mentioned as an important skill gap; this may suggest that workers’ dominance of basic numeracy and literacy are adequate for employers in our sample, or that other skills gaps are more noticeable to employers.

Most Pressing Skills Gaps by Region

As discussed above, though the skills that employers most value did not differ by region, the different education and industry structures may lead to different skills gap profiles in more and less sophisticated economies. The expected trends emerge to some extent where technical skills gaps are most observed by Eastern European employers, while developed country employers note the absence of higher-order
<table>
<thead>
<tr>
<th>Country</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>problem solving</td>
<td>communication</td>
<td>work experience</td>
<td>teamwork</td>
<td>English</td>
<td>Moursheed et al. 2012</td>
</tr>
<tr>
<td>Latin America</td>
<td>critical thinking</td>
<td>communication</td>
<td>life skills</td>
<td>STEM</td>
<td>leadership</td>
<td>Ogier 2009</td>
</tr>
<tr>
<td>Indonesia</td>
<td>English</td>
<td>computer</td>
<td>thinking skills</td>
<td>behavioral</td>
<td>n/r</td>
<td>diGropello 2011</td>
</tr>
<tr>
<td>Lebanon</td>
<td>independent work</td>
<td>computer</td>
<td>Numeracy</td>
<td>foreign language</td>
<td>communication</td>
<td>World Bank 2012</td>
</tr>
<tr>
<td>Macedonia</td>
<td>responsibility</td>
<td>motivation</td>
<td>Communication</td>
<td>customer orientation</td>
<td>literacy</td>
<td>Rutkowski 2010</td>
</tr>
<tr>
<td>Mexico</td>
<td>technical knowledge</td>
<td>leadership</td>
<td>Innovation</td>
<td>teamwork</td>
<td>communication</td>
<td>CIDAC 2014</td>
</tr>
<tr>
<td>Poland</td>
<td>advanced technical skill</td>
<td>responsibility</td>
<td>motivation/commitment</td>
<td>self-management</td>
<td>problem solving</td>
<td>Arnhold et al. 2011</td>
</tr>
<tr>
<td>Romania</td>
<td>professional skills</td>
<td>efficiency</td>
<td>problem solving</td>
<td>customer orientation</td>
<td>planning</td>
<td>Balcar 2012</td>
</tr>
<tr>
<td>Russia</td>
<td>professional skills</td>
<td>problem solving</td>
<td>conscientiousness</td>
<td>decision-making</td>
<td>teamwork</td>
<td>Vasiliev et al. 2013</td>
</tr>
<tr>
<td>Solomon Islands</td>
<td>analytical skills</td>
<td>work attitude</td>
<td>decision making</td>
<td>communication</td>
<td>technical skills</td>
<td>Close 2012</td>
</tr>
<tr>
<td>Tonga</td>
<td>honesty</td>
<td>punctuality</td>
<td>customer awareness</td>
<td>commitment</td>
<td>problem solving</td>
<td>TCCI 2010</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>foreign language</td>
<td>customer awareness</td>
<td>diversity</td>
<td>self-management</td>
<td>problem solving</td>
<td>CBI 2012</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>planning</td>
<td>customer awareness</td>
<td>teamwork</td>
<td>problem solving</td>
<td>communication</td>
<td>UKCES 2012</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>job-specific skills</td>
<td>business awareness</td>
<td>enterprising</td>
<td>vocation-specific skills</td>
<td>customer care</td>
<td>Martin 2008</td>
</tr>
<tr>
<td>United States</td>
<td>language</td>
<td>communication</td>
<td>leadership</td>
<td>work ethic</td>
<td>technical skills</td>
<td>Casner-Lotto and Barrington 2006</td>
</tr>
</tbody>
</table>

*Source: Data drawn from each paper cited in last column.*
cognitive skills. In three (Romania, Russia, and Poland) of the four (Kazakhstan) studies of Eastern European employers (Balcar 2012; Vasiliev et al. 2013; Arnhold et al. 2011), the lack of technical skills was most lamented, followed by socio-emotional skills (efficiency, motivation, and responsibility). In contrast, employers in the US and UK largely identify a lack of higher-order cognitive skills (language, planning), though one of the three UK studies identifies job-specific skills as the primary skill gap. Socio-emotional skills begin to emerge as the second-greatest skills gap in both Eastern Europe and developed country priority rankings. The results are too mixed in East Asia, and the sample is too small in Latin America, Africa, and the Middle East to report regional patterns.

**Most Important Skills Gaps: Managers v. Workers**

Employers in the five studies that disaggregate employer-perceived skills gaps by occupation (UKCES 2012; diGropello 2011; Vasiliev et al. 2013; World Bank 2012; CIDAC 2014) list a higher-order cognitive skill, or socio-emotional skills nine times each (of 25 potential responses) as one of the top five skills gaps faced by their non-manager workers; employers name technical skills gaps four times. The employers emphasize socio-emotional skills as most lacking in their managers, naming them 10 times, compared to seven mentions of higher-order cognitive skills and technical skills gaps, which were mentioned five times.

Although employers value similar skills in managers and workers, the top skills gaps differ by occupation. Similar to the aggregate measurement of the greatest skills gaps, the top-ranked skill gap among managers in three countries are primarily higher-order cognitive skills—decision-making, strategic management, and foreign language. Technical and socio-emotional skills are the most highly ranked skills gap in managers by one country each (Lebanon and Mexico, respectively). In contrast, a range of broad skills sets emerged as most lacking among workers. Employers in two countries named technical skills as most missing in their workers (UK and Mexico) while two others listed socio-emotional skills (conscientiousness and independent work). The fifth country named a higher-order cognitive skill (English language). The second-most important skills gaps for non-managerial workers reveal the same pattern: employers in two countries listing technical skills, and one country each listing higher-order cognitive, basic cognitive, and socio-emotional skills.

**The Skills Development Process and Policy for Developing the Skills Employers Demand**

Employer voices tell us that a broad range of skills are necessary for labor success. Some skills are taught after completing general education, particularly technical
training, while others are commonly acquired at a younger age, namely basic cognitive skills. A relatively large body of literature discusses appropriate methodologies for teaching these skills. However, there is much less guidance on the appropriate ages to learn socio-emotional and higher-order cognitive skills and the appropriate methodologies to teach them. The developmental psychology and education literature provides a framework to guide these skill development processes.

Just as with basic cognitive and technical skills, socio-emotional and higher-order cognitive skills are best taught at specific stages in the life-cycle. Neurological, biological, psychological, and social processes dictate that certain skills cannot be learned before certain ages (Guerra et al. 2014). For example, a toddler is me-centered and is not able to feel the genuine empathy that a primary school student displays. It is not for a lack of being taught to be empathetic, but instead the toddler is not neurologically or psychologically “ready”, and a toddlers’ social context—where she is still largely guided by caregivers—is not conducive to practicing, and thereby developing, this skill. Further, certain skills are the foundation for other skills (Cunha et al. 2005). For example, basic math—which is developmentally appropriate for primary school—is a foundation for secondary-school introduction to physics, just as impulse control is a foundation for the higher-order cognitive skill of problem solving. Heckman (2008) argues that most of the skills gaps at age 18 that help to explain adult outcomes are already present at age five, and that disadvantaged children are at a particular risk of falling behind early and not being able to catch up.

Table 5 presents a rough representation of the appropriate period of the life-cycle during which one may acquire the skills that the employer surveys suggests. In the early years (age 0–5), the most basic cognitive skills can be acquired. Also, some of the most important foundational socio-emotional skills are developed during this period, such as delayed gratification, impulse control, and working with others. During middle childhood (age 6–11), learning really takes off with the ability to rapidly acquire basic cognitive skills—with some higher-order cognitive emerging, such as problem solving—and the child is in a context to develop more complex socio-emotional skills related to engaging and negotiating with others. During adolescence (12–17), the foundations should already be built, the brain is neurologically and psychologically ready, and the social context is appropriate to fully engage in higher-order cognitive development and complex socio-emotional development while still acquiring basic cognitive skills and beginning to develop technical interests and skills. Once reaching early adulthood (18–26), technical skills can be built on the foundation of the basic cognitive skills, higher-order cognitive skills, and socio-emotional skills learned earlier in life.

A lot of skill development occurs outside the classroom, indicating that a wide range of age-relevant actors are best positioned to develop the young person’s skill sets (table 5). Drawing from the Bronfenbrenner ecological risk framework (1979), we see that a young person’s actors of influence broaden, and move away from the

_Cunningham and Villaseñor_
nuclear family, as she ages. In early childhood, family and early childhood development programs are the age-relevant actors due to the biological forces of children being psychologically attached to a core, known family, and to practical issues related to a child’s independence. Thus, these are the actors responsible for developing age-associated skills. During middle childhood, the school gains in importance, as do peers and other mentors, but the family still plays a dominant role. During adolescence, the family starts to fade as peers, educational institutions, and non-family mentors grow in importance, and finally, in young adulthood, socio-emotional skills are refined and shaped by higher education institutions (Robins et al. 2001), work environment, and the family (Roberts, Caspi, and Moffitt 2003). Technical skill development also continues into adulthood through on-the-job training; in fact, once reaching adulthood, firms are the primary source of new skills acquisition for workers (Villaseñor 2013).

Table 5. Skill Formation at Different Points of the Life-cycle

<table>
<thead>
<tr>
<th>Period</th>
<th>Skill type</th>
<th>Age-relevant Actors</th>
<th>Sample programs to Guide Actors to Build the Skills (for a list of evidence-based programs, see Guerra et al. 2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Early years</strong></td>
<td>Basic cognitive</td>
<td>Family, ECD programs</td>
<td>Quality parenting (Nuevo Postnatal, Program on Cognitive Development, Early Enrichment Program)</td>
</tr>
<tr>
<td>(0–5)</td>
<td>Foundational</td>
<td></td>
<td>Child-focused ECD (Perry Program, Head Start)</td>
</tr>
<tr>
<td></td>
<td>Socio-emotional</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Childhood</strong></td>
<td>Basic cognitive</td>
<td>Family, schools, peers</td>
<td>Holistic curriculum, teaching methodology, and monitoring and evaluation system (KIPP, EPSIS, Enseña Chile, RCCP)</td>
</tr>
<tr>
<td>(5–12)</td>
<td>Socio-emotional</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Foundational</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Higher-order cognitive</td>
<td>Schools, peers, mentors, family</td>
<td>After-school/extra-school/activities (BBBS, Student Success Teams)</td>
</tr>
<tr>
<td></td>
<td>Initial technical</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Adolescence</strong></td>
<td>Basic cognitive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(13–17)</td>
<td>Socio-emotional</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Higher-order cognitive</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Technical</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Early (17–29) and Middle (30+)</strong></td>
<td>Socio-emotional</td>
<td>Higher education institutions, training institutes, work place, family</td>
<td>Apprenticeships (Jóvenes programs)</td>
</tr>
<tr>
<td><strong>Adulthood</strong></td>
<td>Higher-order cognitive</td>
<td></td>
<td>Experiential on-the-job training Skills certification system, support systems for worker transition to firms where new learning can occur</td>
</tr>
<tr>
<td></td>
<td>Technical</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Own elaboration based on Banerji et al. (2010) and Guerra, Modecki, and Cunningham (2014).
There are a multitude of methods for effectively teaching the appropriate skills by each actor at each life-cycle stage (table 5). For parents of young children, good family leave policies that allow parents to provide quality parenting and programs to enhance parental learning and encouragement of early stimulation and nutrition have been shown to foster greater acquisition of cognitive skills and socio-emotional skills in a child (Gertler et al. 2014; Kagitcibasi 1988). Child-centered ECD that focuses on improving personality traits and managing externalizing behaviors while also acquiring basic cognitive skills have shown positive results in employment, wages, and positive behaviors for more than 30 years after program participation (Schweinhart et al. 2005).

Programs to help develop socio-emotional learning in middle childhood are largely designed for the school setting, and are complemented by parental engagement. CASEL (2013) recommends four prerequisites for developing socio-emotional skills in schools: emotionally supportive teachers, a positive school climate, socio-emotional learning integrated into subject-based lessons, and free-standing lessons to teach socio-emotional skills. These may take place in the classroom via regular curriculum and instruction activities, or through complementary programs. However, the research finds that the most successful socio-emotional learning occurs when incorporated into the day-to-day curriculum and when connected to other school activities (Greenberg et al. 2003). For example, the US Knowledge is Power Program (KIPP) implemented in primary and secondary schools sets expectations, requires behaviors grounded in good socio-economic skills, and works closely with each child and their families to ensure success (Angrist et al. 2010). Similar programs exist in developing countries as well (Heckman and Kautz 2012; Alfonso et al. 2012). Impact evidence shows these methods positively affect socio-emotional skill development, associated positive behaviors, and cognitive learning outcomes, especially in middle-childhood (CASEL 2013; 2015).

Both middle childhood and adolescence are key periods for developing higher-order cognitive skills, and pedagogical methods can be powerful tools for fostering these skills. The most effective methodology for teaching problem-solving, creativity, communications, and overall thinking skills is largely through student-centered, structured pedagogical methods (King et al. 1998) where the educator uses questions, guided practice, team activities, student discussions, and peer tutoring, for example, to engage students to solve problems, test their ideas with other students, and express themselves verbally and in writing (Brookhart 2010). When the same material is taught through traditional didactic methods to a control group and interactive methods to a treatment group, the group receiving the interactive teaching has a statistically significant greater performance on language and math exams than the group receiving the didactic instruction (Brookhart 2010).29

While adolescents benefit from the school-based programs (CASEL 2015), their greater independence from the family sphere and initial transition to world of work implies that mentoring programs and technical training can also contribute to their...
socio-emotional and higher-order cognitive skills development process. Mentoring programs can take different forms such as after-school clubs, programs that pair model adults with children, or sports programs run by child development specialists; the former two models have shown to increase cognitive and socio-emotional skills of participants relative to control groups (Tierney and Baldwin 2000; Boys & Girls Clubs of America 2004). As the school-to-work transition process begins, job training and apprenticeship programs should be augmented by socio-emotional and higher-order cognitive skills development. For example, such a program in the Dominican Republic has been shown to increase leadership, conflict management, self-esteem, interpersonal skills (for some groups), organization skills, empathy, and hard work, in comparison to a randomly selected control group that did not receive the intervention (Vezza et al. 2014; Ibarraran et al. 2012).

Skills development in early adulthood is primarily focused on acquiring technical and higher-order cognitive skills; programs for socio-emotional skill development are quite rare. There is an increasing consensus that technical training models need to reform to better teach the range of skills that employers demand (OECD 2014; Fawcett 2014). Specifically, training should be delivered through structured experience-based (rather than classroom) learning that reinforces all four broad skills sets by real-world practice in a supportive context. Although experiential learning is common practice in European technical training systems (Fawcett et al. 2014), employers in other countries can find room for improvement. For example, Latin American executives identify “work-study programs that bring students the workplace” as the primary way that firms can contribute to skills development, though few engage in the practice (Ogier 2009). The public sector can provide technical assistance and financial incentives to alleviate the costs of designing and implementing a training program while not affecting firm profitability. For example, Colombian law (Law 789 of 2002) provides tax breaks to firms to train their workers. Linked to this is the emerging recognition that instructors should be recruited from the private sector who can provide hands-on knowledge. Most Latin American executives in a small sample thought that the lack of teaching skill to teach in an applied manner is a bigger problem than educational infrastructure, curriculum, or teacher qualifications in the subject matter (Ogier 2009).

Conclusions

The review confirms that there is a mismatch between the education sector’s perception of skills demand and that of the productive sector. While the education sector believes that it prepares students well for the labor market (Moursesh, Farrell, and Barton 2012; IFC 2011) by focusing on basic cognitive and technical skills development, a review of 24 studies that measure employer skills preferences...
and 15 studies that measure key skill gaps find that employers have a different view of the most important skills for labor market success and where the principle skills gaps are.

While employers around the world value all four skills sets, socio-emotional and higher-order cognitive skills are consistently prioritized. More than 50% of the top five skills identified by employers can be classified as socio-emotional skills, and 79% of the top-ranked priority skill are a socio-emotional skill. Another 30% of the top five skills can be classified as higher-order cognitive, with a top ranking in about 17% of the cases. Oral communication—a higher-order cognitive skill—ranks consistently very high, as do a small set of socio-emotional skills, namely ethics, punctuality, honesty, and teamwork.

Technical skills, which are often assumed by policymakers to be equivalent to job skills, were ranked as third-most important in the aggregate estimate, but they emerge strongly for some groups. Specifically, Eastern European employers add technical skills to the list of priority skills sets, joining socio-emotional and higher-order cognitive skills. Employers in all other regions and in the US do not value technical skills as highly. This variable was difficult to analyze, though, since the classification of “technical skills” may have over-simplified employer preferences since all “job-related” responses were put in this skills set, even though many job related skills are socio-emotional or higher-order cognitive by nature. Technical skills seem to be complements to, not substitutes for, cognitive and socio-emotional skills.

Basic cognitive skills were rarely prioritized, which may reflect that these skills are not needed or, more likely, that they are in sufficient supply that employers do not notice how important they are. This interpretation is supported by the lack of mention of basic cognitive skills among managers, who most certainly need these skills to perform their jobs well.

The aggregate patterns hold up when dividing the sample by region, industry (manufacturing v. service firms), occupation (managers v. workers), and education level of the workforce; socio-emotional and higher-order cognitive skills emerge most strongly for each sub-group. And while the priority skills sets for managers and workers or for less and more educated workers are the same across countries, priority skills sets for manufacturing versus service sector differ across the three countries in the sub-sample. These results should be understood with a large margin of error given the very small size of the samples.

Employers identify significant skills gaps. The aggregate analysis shows that socio-emotional skills were most cited among the top five skills gaps, but higher-order cognitive skills were the top-ranked skill gap. About 42% of the top five skills gaps can be classified as socio-emotional skills. However, they dominate the second through fourth priority rankings. Instead, higher-order cognitive skills are identified as the most important types of skills in nearly half of the sample. When
disaggregating by region and occupation, slightly different trends emerge. Among regions, technical skills are named as the priority skill gap in Eastern Europe, but not named as the top skill gap by any study in any other region. Socio-emotional skills are particularly important in developing regions compared to the developed countries in the sample.

Although socio-emotional and higher-order cognitive skills are consistently valued by employers, they often fall outside of school curriculum or teaching methods. When we bring employer preferences together with the skills formation process as understood by developmental psychologists, four key conclusions for education/skills development policy emerge. First, the skills development process necessarily begins at birth (or before) and continues throughout the life cycle. Certain skills that employers demand are formed in the toddler years and other skills can only be developed once the foundational skills are there. Waiting until school is completed to begin developing job-relevant skills is too late. Second, schools can play a much larger role in socio-emotional skill development by developing a teacher’s personal skills, strengthening the school climate so it encourages positive behaviors, weaving socio-emotional skills development and practice into pedagogical methods, and teaching these skills as part of the curriculum. Third, other actors should play a larger and more structured role in skills development. Certain skills are better taught by parents, mentors, or colleagues in the work place. This points to an education/skills development strategy and related programs to support the actors that are best suited to provide instruction to children at each age-appropriate stage. Fourth, the skills most demanded by employers—higher-order cognitive skills and socio-emotional skills—are largely taught (the former) or refined during adolescence, which argues for a secondary school general education until these skills are formed. Rather than the early tracking of youth into technical training (ranked third by employers), skills/education systems need to ensure that the foundational basic and higher-order cognitive and socio-emotional skills are there to allow for effective technical skill acquisition.

While this paper presents evidence from the literature, more work needs to be done to definitively guide policy. First, the research needs to move beyond simple correlations and into marginal effects, namely, to generate parametric estimates that will better define the skills sets demanded by employers in different industries and occupations. Second, the question of revealed versus stated preferences merits greater research to determine the validity of employers preference surveys. While this paper presents evidence that these preference sets are very similar, by comparing employer survey skills to those used by for-profit headhunting firms and want ads, a direct analysis is necessary to understand the extent to which this gap may bias employer survey findings. Finally, the absence of consistency in definitions and measurement makes cross-country comparisons difficult. The research needs to move toward a common set of concepts to allow for cross-country learning.
Many thanks to Monica Parra and Mariana Escalante for their contributions to the paper, and to Stephen Close, Margaret Grosh, Katia Herrera, William Maloney, Noel Muller, Jan Rutkowski, Fred Oswald, and Alexandria Valerio for their critical review and suggestions.

Notes

Wendy V. Cunningham is a Lead Economist at The World Bank; wcunningham@worldbank.org. Paula Villaseñor is the Social Promotion and Participation Coordinator for Upper Secondary Education at the Mexico Ministry of Education; paulavillasenor@gmail.com.

1. Hanushek and Woessman (2008) identify two sources of measurement error. First, there is substantial heterogeneity in the skills acquired at each level of schooling across countries, regions within countries, and schools within regions. Second, much skill acquisition occurs outside of the classroom (Hanusheck 1979).

2. See Almlund et al. (2011) and Borghans et al. (2008) for reviews of this literature.

3. Autor, Levy, and Murnane (2003) illustrate this phenomenon by arguing that as firm technology improves, repetitive and predictable tasks are automated and workers performing routine tasks are substituted by computers, which themselves need to be complemented by workers who perform non-routine problem solving.

4. We intentionally include a broad geographical mix of countries to permit us to identify if there are global trends in skills demand or if skills demand and shortages differ by region or country.

5. Although the economics literature divides skills into cognitive and non-cognitive skills, psychologists argue that the latter is a misnomer since they clearly involve cognition. With a nod to our psychology colleagues, this paper instead uses the term “socio-emotional skills.” For a light-hearted review of terminology, see http://www.npr.org/sections/ed/2015/05/28/404684712/non-academic-skills-are-key-to-success-but-what-should-we-call-them?utm_campaign=storyshare&utm_source=twitter.com&utm_medium=social.

6. The concept of “higher” and “lower” order (or, what we call “basic”) cognitive skills emerges from Bloom’s taxonomy of learning, where knowledge and comprehension are classified as “lower order” skills; built upon these are the higher-order skills encompassed in critical thinking (Bloom et al. 1956).

7. This study finds that the impact of school attainment on wages falls from 7.1–5.9% after controlling for literacy scores.

8. The psychology literature defines technical skills as a sub-set of cognitive skills (Almlund et al. 2011).

9. Tan and Nam (2012) review recent studies estimating the wage premium for technical training to general education and find higher returns to the latter. The results are merely suggestive, though, since the reviewed studies do not control for unobservable skills, which may be both correlated with higher wages, and with the choice of education path, thus potentially biasing the comparison.

10. The idea that socio-emotional skills are an important driver of economic success can be traced to Bowles and Gintis (1976), who explain that a measurable part of the variance in earnings among observationally equal individuals, particularly those with equal levels of education, are due to behavioral skills.

11. The socio-emotional skills that are most commonly measured and correlated in the economics literature are limited to personality trait facets based on the Goldberg Big 5 construct (Goldberg 1993), listed above. While this construct has been validated across cultures and time, it is limited in that it does not measure behaviors or attitudes.

12. See Almlund (2011) for a review of literature.
13. This search produced papers from literature in economics, human resources psychology, industrial sociology, management, and education.

14. Research Papers in Economics (IDEAS-RePEc) is the largest bibliographic database dedicated to Economics in which most of the papers are fully downloadable; see http://ideas.repec.org/. We searched the following JEL codes: J23 (Labor Demand) and J24 (Human Capital; Skills; Occupational Choice; Labor Productivity).

15. EconLit is the American Economic Association’s electronic bibliography; see www.aeaweb.org/econlit/.

16. See www.eric.ed.gov; ERIC is one of the primary database for education literature.

17. The website www.psycontent.com is the database of the American Psychological Association, specialized in psychology and psychiatric journals, and includes the following databases: PsyJOURNALS, PsycARTICLES®, PsycINFO® and PsyCOLLECTION®.

18. The earlier literature is summarized in Cotton (1993). We include one paper from 1997 in order to balance out the samples that use US data. The recent papers only ask employers about skills demand for more educated workers; Zemsky (1997) provides information about the median worker.

19. A detailed description of each study sample frame is available from the authors.

20. Anecdotes (from the authors’ interviews) may illustrate the point. An owner of a bottling company in Tonga said that when interviewing, he first considers the body language when meeting the candidate, then assesses eye contact, and, once the person passes that phase, he asks about technical skills. A psychologist working with youth in Jamaica tells of a young woman who, after several failed interviews, asked if the profane tattoo on her forehead may have been a factor in her lack of success.

21. Many surveys do not go into detail on technical skills, so “job-relevant skills”, computer skills, and “work experience” are included in this set.

22. The socio-emotional skills defined in the employer surveys differ from those in the economics literature. The standard research uses umbrella concepts that encompass a range of attitudes and behaviors while employers indicate very specific attitudes and behaviors. For example, the “teamwork” skill that many employers define is not included in the standard literature, but teamwork concept is encompassed in the extraversion and agreeableness umbrella concepts that have been well studied. Guerra, Modecki, and Cunningham (2014) map the 140 employer skills identified in this paper to the Goldberg Big 5 construct most commonly used in the economics literature.

23. Several studies only allow employers to priority rank the top five skills or skill gaps. Thus, we also limit r to five.

24. The data do not provide sufficient information to disaggregate skills gaps patterns by industry or by education level of the worker.

25. Ability to collaborate and cooperate with others, control emotions, and avoid negative reactions

26. The data in this review do not allow us to estimate partial correlations. However, the Indonesia study provides some insight into the question of partial versus simple correlations since it observes the prioritization of two dimensions simultaneously: industry and occupation.

27. Unlike in the sample which asks about the most important skills, the data on key skill gaps is disaggregated across skill level and type of industry in many of the studies. Since we do not know the share of the labor force in each of these categories, we cannot appropriately weight the responses. Thus, we make a very general assumption that the labor force is equally distributed across these categories and we take a simple average across skill level or industry.

28. Contrary to decades of assumptions of the non-malleability of personality traits, psychologists are finding that personality change continues into adulthood once careers have been established that shape personality more profoundly than transitory early adulthood jobs (Roberts 1997), and as life changes such as marriage affect personality (Robins, Caspi, and Moffitt 2002).

29. Given the difficulty of directly measuring higher-order cognitive skills (see King et al. 1998 for a discussion), they are indirectly measured through performance on content tests that require the use of higher-order cognitive skills to score higher.
References

*Indicates the reference is a data point for this paper.


130


Revisiting the “Cash versus Food” Debate: New Evidence for an Old Puzzle?

Ugo Gentilini

The longstanding “cash versus food” debate has received renewed attention in both research and practice. This paper reviews key issues shaping the debate and presents new evidence from randomized and quasi-experimental evaluations that deliberately compare cash and in-kind food transfers in ten developing countries. Findings show that relative effectiveness cannot be generalized: although some differences emerge in terms of food consumption and dietary diversity, average impacts tend to depend on context, specific objectives, their measurement, and program design. Costs for cash transfers and vouchers tend to be significantly lower relative to in-kind food. Yet the consistency and robustness of methods for efficiency analyses varies greatly. Social protection, Safety nets, Food security, Political economy, Cash transfers, In-kind transfers, Food aid, Vouchers, Food stamps, Impact evaluations, Cost-effectiveness. JEL codes: D610, H530, O120, O130, Q180

Introduction

There is little doubt that cash transfers are amongst the most rigorously-evaluated interventions in development. The basic question that this paper addresses, however, is not whether cash transfers work in general, but whether and why they do so relative to in-kind food assistance. Indeed, the cash versus food debate is among the most longstanding, controversial, and polarizing social protection quandaries. Consider the following quotes:

[T]he big reason poor people are poor is because they don’t have enough money, and it shouldn’t come as a huge surprise that giving them money is a great way to reduce that problem — considerably more cost-effectively than paternalism [i.e. vouchers and in-kind transfers]. So let’s abandon
the huge welfare bureaucracy and just give money to those we should help out. (Kenny 2013)

[T]housands of economics professors across the globe persuade millions of undergraduates that granting the poor distribution of benefits in-kind is less “efficient” than simply transferring to them cash (...). The economist’s traditional, normative dictum on benefits in-kind may be analytically elegant (...) but practically dead wrong. (Reinhardt 2013)

These excerpts seem to posit two irreconcilable perspectives. But is there some truth in both? What does the cross-country evidence and experience tell us about transfers’ relative performance? Until recently, a key limiting factor in informing the debate has been the paucity of robust comparative evidence. As a result, the discussion on transfer selection has been largely based on inference—that is, based on findings from programs implemented in diverse contexts, relying on different evaluation methods, or designed for differing purposes. Such extrapolation and comparison of findings could, at best, be suggestive of relative effectiveness.

This paper reviews key findings emerging from a new generation of robust impact evaluations in ten developing countries, namely Bangladesh, Cambodia, the Democratic Republic of Congo, Ecuador, Ethiopia, Mexico, Niger, Sri Lanka, Uganda, and Yemen. These studies compare cash and food transfers under the same circumstances through randomized controlled trials (RCTs), quasi-experimental methods, or regression analysis. The paper, which frames and discusses emerging results within the broader issues that shape transfer selection debates, argues that the debate involves a range of factors that makes it more complex than often assumed. Moreover, it shows that, in terms of effectiveness, there seems to be no systematic dominance of one transfer over the other, and that a number of information gaps persist. Overall, the paper argues that transfer selection may deserve further attention on a par with other program design features such as choices around conditionality and targeting methods.

The analysis is conducted within certain parameters. First, the article examines the transfer debate within the universe of safety net programs, including interventions such as conditional cash transfers, school feeding, public works, or social pensions. In those programs, the size of transfers is generally devised to enable access to food and meet some essential needs. According to survey data from 69 countries, the average size of safety net transfers is about 23 percent of the poor’s income or consumption (World Bank 2014). Therefore, the vast body of literature on “cash grants”—which involve relatively large sums of money to jump-start entrepreneurship and investment—go beyond the scope of this paper. Similarly, in-kind transfers may encompass a vast array of items such as shelter, agricultural inputs, and subsidized health care. Yet we only consider a subset of those transfers, namely food commodities; as such, we use the terms “in-kind” and “food” interchangeably. Also, we
regard food as a tout-court transfer, while de facto food rations often include a bundle of commodities of various monetary and nutritional values.

The reminder of the paper is organized as follows: the next section sets out a framework examining key areas that influence transfer selection debates. The subsequent section describes the compilation of evaluations and sets out their results. The next section provides a discussion of the implications from the analysis, while the last section concludes.

Anatomy of the Debate

The issue of transfer selection is relevant under a range of poverty-reduction debates, but it is particularly compelling for issues around food security. The role of social protection in food security is being increasingly examined, including through the lens of its three core pillars: how social protection can enhance food production and agricultural productivity (Tirivayi et al. 2013); how transfers can provide access to food (CFS 2012); and how social protection can, as Alderman (2014) put it, be more nutrition-sensitive. A fourth food security pillar, risk or stability, underpins those three dimensions and it is central to the connections between social protection and resilience (Davies et al. 2013). The interactions between social protection transfers and the abovementioned food security components are mediated by a range of other factors (e.g., food security being only one factor in the nutrition equation), and occur under different timeframes (e.g., the likely timeframe for reducing malnutrition is measured more in terms of a generation, rather than of a few years).

Three forms of transfers are considered in this paper, namely cash, food, and vouchers. Cash transfers provide people with money, while food transfers (or food aid) include the distribution of commodities. Food transfers can be procured internationally (imported food) or local-sourced in the same country where the program is implemented. Vouchers are also known as stamps or near-cash transfers and can be used in predetermined locations, including selected shops, supermarkets, retail stores and fairs. Vouchers take two forms: on one hand, “value-based” vouchers provide access to commodities for a given monetary amount; on the other hand, vouchers can be “commodity-based”, or tied to a pre-defined quantity of given foods. Therefore, vouchers are a hybrid form of transfer that display features of both cash (value-based vouchers allow for some level of choice, although this is limited to the commodities available in the chosen location) and in-kind food (the implementation of commodity-based vouchers can closely resemble that of public food distributions systems).³ Cash transfers can be considered the modality that provides beneficiaries (consumers) with largest choice while at the same time stimulating local markets. Also, vouchers can equally bolster local (food) markets, although the
choice they provide is more restricted or none. Food transfers provide no choice, but can stimulate markets if procured locally, although at a different level than vouchers: indeed, the former tend to entail the engagement of a more limited number of upstream, larger producers, while the latter would directly involve a larger number of downstream, smaller actors. From this perspective, vouchers are more “local” in nature. Against this background, the reminder of the section briefly discusses a set of issues in the spheres of theory, political economy, and technical matters that will be useful when discussing the evaluations in the following section.

Theory and Political Economy

The microeconomic foundations of the debate tend to draw from neoclassical theory as laid out by Southworth (1945). The transfer debate received considerable empirical attention in the 1980s and mid-1990s, including the production of a vast body of literature examining the United States’ Supplemental Nutrition Assistance Program (SNAP; Basu 1996; Faminow 1995; Fraker et al. 1995; Devaney and Moffitt 1991; Levedahl 1991; Coate 1989; Blackorby and Donaldson 1988). Interestingly, an empirical regularity in the literature seems to be the presence of a “cash-out puzzle” rejecting the neoclassical hypothesis—that is, “inframarginal” cash and food transfers should hypothetically have an equal effect on food consumption. Indeed, “... virtually every study finds food stamps increase household nutrient availability at 2 to 10 times the rate of a like value of cash income,” (Barrett 2002). While there seems to be an overall difference in outcomes between cash and in-kind transfers, there remains great dispute as to why such difference persists.

The transfer debate is not new. Historians found accounts of the quandary in 17th Century England, and even in ancient Edessa around 124 BC (Lindert 2005; Garnsey 1988). More recently, from the late-1990s and mid-2000s, the emergence of cash transfers on the global stage was in part defined by mounting criticism over the international food aid system, particularly in terms of transoceanic shipments and local monetization of commodities (Barrett and Maxwell 2006, 2005; Clay 2006; OECD 2005). Since the mid-2000s, the growing application of technology to the delivery has further propelled interest in cash as an efficient way of providing assistance (Vincent and Cull 2011).

While the global food aid model has evolved dramatically toward more targeted, technology-supported, locally-procured and nutritionally-sound commodities, some lingering perceptions around food transfers seem hard to fully dispel. In part, the mixed performance of large-scale national public food distribution systems may have contributed to a view of food as an intrinsically second-best modality (Gol 2005). A related political dimension is that an appearance of a linkage to food may
be attractive not only to taxpayers, but also with food producers and possible lobbies.

Also, in some contexts food transfers may still evoke the image of long truck convoys pouring onto famine-stricken areas, or of forces behind food distributions driven by commercial and geopolitical interests. More subtly, in-kind food may be somewhat more directly associated with the inability of countries or households to feed themselves, hence swiftly juxtaposing the issue of food provision with that of personal dignity. Instead, cash transfers tend to conjure just the opposite—images of freedom and independence—even if they are supposed to play the exact same role as food transfers (e.g., providing access to food).

The latter observation raises perhaps the most fundamental philosophical dilemma in the debate, namely the trade-off between providing choice and ensuring (as much as possible) a positive externality. On one hand, a recurrent view of in-kind transfers considers them as inherently paternalistic. From such a standpoint, the arguments in favor of in-kind transfers revolve around an “over-provision” of goods (i.e., extramarginal transfers). This involves informational, principal-agent, or behavioral arguments that often provide the foundations for much of the debate around conditionalities (Fiszbein and Schady 2009). In broad terms, this suggests that the expected externalities of an in-kind transfer would be desirable when there is a private under-investment—say, in food consumption or nutrition—below an optimal social (or even private) level. On the other hand, the provision of in-kind transfers represents a vehicle to influence behaviors (e.g., consumption patterns). To some extent, paternalism can less judgmentally be considered an approach to address externalities in health, or the fact that parents make decisions for children (thus their welfare may be discounted), or the parents are misinformed about nutritional benefits (Das et al. 2005). Under this view, changing consumption via prices or modified preferences is a motivation for in-kind programs.\(^ {11} \) Shifting gender control is a subset of this motivation. On the other hand, in-kind transfers do not allow people to maximize their utility through choice, that is, the magic of cash lies in its fungibility. Also, cash helps shift the balance of power. As Devarajan (2013) states, “... cash transfers have the potential to shift not just poverty-reducing policies but also the balance of power between government and its citizens, in favor of the latter.”

In other words, the mainstream argument in favor of cash is straightforward: cash provides choice and transfers power, hence making them “redefinitional”.\(^ {12} \) This is an undisputable advantage, though it needs some qualifications: can the provision of choice be considered an objective per se? In that case, any cash transfer can be considered “effective” and would not require evaluations. Or would the same effect be achieved through in-kind transfers? An inframarginal in-kind transfer has the same income-transfer effect. As such it can free up disposable income. And can choice be exercised effectively when there are limited supplies of goods or they are...
unaffordable? This is intimately linked to technical issues around market functionality, which will be discussed in the next section. Finally, an enlarged definition of “choice” may include the ability to choose what to receive in the first place. If the poor know best what they prefer, then they should be provided with choice on what to get (e.g., cash or food), not only on how to use a given transfer (i.e., cash). For example, in 2006 a combined cash and food transfer program was implemented in rural Malawi. The follow-up evaluation explored participants’ preferences over one or both transfers and found that “…most beneficiaries were very satisfied with receiving both food and cash.” (Devereux 2008). Yet the redesign of the project in 2007 included only cash transfers. The point here is the need for a broader notion of “paternalism”—that is, one that is not associated with a given transfer modality, but instead based on the extent to which interventions involve and engage beneficiaries as key stakeholders throughout the life of the program.

Ultimately, the tension between fungibility and paternalism seems ingrained in political economy and philosophical debates that go beyond the choice of transfers (Pritchett 2012, 2005; Lopez-Rodriguez 2011; Currie and Gahvari 2008). These debates nevertheless affect the selection of transfers, including through factors such as the contextual definition of what poor people “deserve”, the preferences expressed in a given polity and social contract, public perceptions as rooted in historical precedents, the influence of particular constituencies, and to some extent even the symbolic value carried by food in societies.

Factors Affecting Choices

The available transfer literature can be grouped, in addition to the earlier SNAP studies, into five broad “empirical buckets”. First, there is high-quality research conducted on individual cash-based programs (Evans et al. 2014; Baird et al. 2013; de Brauw et al. 2012; Barham and Maluccio 2009; Duflo 2003). Second, there are comprehensive reviews focusing on cash transfers, often in relation to sectors, themes, and regions (Alderman 2014; Davis 2013; Garcia and Moore 2012; Bailey and Harvey 2011; DFID 2011; Fiszbein and Schady 2009; Adato and Bassett 2008). Those two threads of literature are mirrored on the food front: in-kind food transfers have been carefully examined empirically (Kazianga et al. 2014; Decon et al. 2012; Adelman et al. 2008a, 2008b; Gilligan and Hoddinott 2007; Yamano et al. 2005; Del Ninno and Dorosh 2003; Quisumbing 2003), as well as being subject of cross-country reviews and meta-analyses (Alderman and Bundy 2012; Webb et al. 2011; Del Ninno et al. 2007; Rogers and Coates 2002). Finally, publications have documented and summarized evidence emerging from all those four sets of studies (Lentz al. 2013; Gentilini 2007).

While recalling the general caveats on comparability, those studies have generally concluded that transfer appropriateness is context-specific and hinges on multiple
factors. These include program objectives, the functioning of markets, administrative capacity, seasonality, security, intra-household preferences, and community dynamics. These factors largely shape the performance of transfers across time and space.

Turning “needs” into “effective demand” is a key rationale for cash transfers. Yet this might be challenging in the presence of weakly integrated and competitive markets. In those contexts, price transmissions would not necessarily signal relative scarcities, and localized cash injections may result in price spikes leaving consumers or net buyers worse off.\(^{13}\) Ensuring that markets would respond to an increase in effective demand is key to ensure that purchasing power is maintained and local multipliers enabled. From this perspective, a basic level of market functioning is a prerequisite for the effective provision of cash transfers.\(^{14}\)

The discussion on market analysis has important practical implications for program design, implementation, and efficiency. In some cases, prices may be particularly volatile, with a certain degree of unpredictability in future trajectories. These could turn a program that was efficient in the design stage into a cost-inefficient program during implementation. Indeed, keeping purchasing power constant in the wake of sharp and protracted price increases may escalate costs due to extensive use of contingency funds, such as shown in Zambia (Harvey and Savage 2006).

Price dynamics also have an influence on beneficiaries’ preferences, which tend to lean towards in-kind transfers when prices are higher.\(^{15}\) This was clearly documented in the case of Ethiopia during the sudden increase in wheat prices in 2008 (Sabates-Wheeler and Devereux 2012). While it is difficult to generalize people’s preferences for a certain transfer modality, some general patterns can be discerned. As a natural consequence of price fluctuations, the preference for cash, vouchers, or in-kind food aid can vary by season. Yet gender also matters, and it is a generally observed pattern that women spend resources differently from men (Doss 2013). In a number of societies, however, women tend to prefer food, which they are more likely to control, while men may prefer cash transfers. However, anecdotal evidence shows that beyond the fact of who receives the transfer, the process of intra-household decision making also counts (e.g., men and women deciding in concert how to use household resources, even in contexts where women may be constrained in reaching markets, or face risks to access them, such as in refugee camps). As such, intra-household and gender-specific preferences are more complex and dynamic than often assumed. For instance, very few studies support the unitary model of household behavior, and even fewer, if any, are experimenting with the differential impacts (e.g., on nutrition) of cash versus food transfers as provided to men and women (Braido et al. 2012).

New research is also shedding light on the intra-community effects of transfers, some of which are positive and empowering, while others generate undesired...
consequences in terms of social relations. For example, in Zimbabwe it was observed that, unlike food, cash transfers were not shared within the community, thus hindering informal mutual support and risk management mechanisms among members (MacAuslan and Riemenschneider 2011).

Returning to the issue of price trends and real value, predictable price seasonality has been recognized as one of the key drivers of food insecurity in Africa. For example, in Malawi analysis on price trends over 20 years show mean inter-seasonal price fluctuations in the order of 60 percent (Ellis and Manda 2012). There are various emerging options to deal with marked seasonal price fluctuations; one way is index-linking cash transfers to food prices, such as in the Dowa Emergency Cash Transfer program in Malawi. In that context, the approach “... protected household food security until prices started falling just before the next harvest, [although] it required a degree of administrative and budgetary flexibility that is inconceivable for most governments and donor agencies,” (Devereux 2012). Another option is to deliver transfers half in cash and half in food. For example, in Swaziland a program was designed in 2007/08 where people were given a half ration of food (maize, beans, and oil) plus the equivalent in cash, each month for 6 months from November 2007 until the harvest of April 2008 (Devereux and Jere 2008). A third alternative is the distribution of cash and food transfers by season, with food provided in the lean season and cash immediately after the harvest. That approach, for example, is implemented under the Productive Safety Net Program (PSNP) in Ethiopia. This is likely to have a number of advantages, although it may also entail considerable analytical planning and logistical coordination among modalities.

More generally, cost-efficiency is one of the key factors that favor adopting cash-based transfers. Studies assessing relative costs showed that they tend to be higher for food transfers than for cash transfers and vouchers. This is not surprising, given the logistics that food programs normally entail (e.g., transport, storage, handling, etc.). Yet cost assessments seem not to adopt the same standard of rigor as impact assessments. Indeed, there are limited comprehensive protocols on cost analysis, with comparisons often being hampered by the lack of consistency and clarity around data quality, estimation strategies, and cost structures. For example, a review of 27 programs showed that administrative costs for cash transfers and vouchers vary from 3 to 37 percent of total cost (Meyer 2007). Clearly, such a range may have as much to do with different formulas and accounting methods as real differences in cost. Furthermore, a frequent practice in the literature is to convert the cost of food commodities based on market prices at the retail level (upon which the size of cash transfers are often determined). This may significantly overestimate food costs, as they ignore the economies of scale from procuring large quantities of commodities at the producer-level. Preliminary calculations for a sample of countries have shown that those cost differentials could even offset the larger logistics costs of food programs, making in-kind transfers surprisingly more efficient.
than cash transfer programs. For example, Creti (2011) estimated that in the West Bank and Gaza, the costs for vouchers would be from 2 to 2.5 times higher than for in-kind food commodities of similar energy and nutritional content. Similar results emerged for a voucher program in Kenya, but not in Bolivia (ACF 2012).

Impact Evaluations

This section presents new evidence emerging from a set of quantitative evaluations, most of which are taken from randomized and quasi-experimental studies. The next two subsections describe those studies and their results.

Descriptive Features

We examine 11 impact evaluations in 10 countries published over the period 2006–2013. These include the complete gamut of comparative experimental and quasi-experimental trials that, to our knowledge at the time of writing this paper, have been conducted on the matter in developing countries. The studies were fielded in various contexts and include responses to sudden natural disasters (Sri Lanka), slow onset crises (Niger) and man-made emergencies (Congo, Ecuador); others are implemented as part of long-term, institutionalized social protection systems (Bangladesh, Cambodia, Ethiopia and Mexico), or envisage developmental interventions in fragile contexts and regions (Yemen, Uganda). Public works were included in 3 cases, while conditional and unconditional programs were present in 4 and 7 countries, respectively (4 countries had combined interventions).

Some notable differences between countries emerge. For example, the size of monthly transfers range from $5 in Cambodia to $50 in Niger; similarly, the share of transfers in household expenditures varies from 2.5 percent (Cambodia) to 30 percent (Bangladesh). The composition of food rations also vary, ranging from one commodity (rice) in Cambodia to up to 11 commodities in Mexico. Programs could have durations as different as a one-off 3 months intervention in Sri Lanka to a full 12 months (for 4 consecutive years) in Bangladesh. Distribution frequency and delivery mechanisms also differ, while evaluation methods include RCTs, quasi-experimental techniques (difference-in-difference, propensity score matching), regressions, and combinations thereof.

In particular, the Programa de Apoyo Alimentario (PAL) in Mexico provides unconditional food transfers to most of the target households, while at the time of data collection (2004–2005), cash transfers were provided to approximately 5 percent of beneficiaries living in villages so isolated that PROGRESA could not be implemented there. The transfer was of 150 pesos monthly, equivalent to $13, which is
equal to the cost of the food basket in local markets. No adjustments to transfer amounts were made for family size or composition. The transfer is delivered bi-monthly through stores in the case of food, and biometric debit cards for cash transfers. The PAL experiment was carried out in 208 villages that were randomized into four treatment groups: an in-kind transfer plus educational classes (the standard PAL treatment); in-kind transfer without classes; a pure cash transfer of 150 pesos per month plus classes; and finally a control group. Cunha (2014) uses the experimental PAL data to estimate the impacts between the cash and food transfers using a difference-in-difference estimation strategy. Moreover, we completed Cunha’s study by examining the previous evaluation by Skoufias et al. (2008). Using the same estimation method and sample size, their analysis offered additional insights for poverty and labor market participation.

In Niger, a large-scale public works program was implemented in 2011 over a six-month period, from April to September. The program included public works and unconditional transfers implemented in sequence for the same group of beneficiaries. In some villages, registered households received cash, while in others they received food. In cash-receiving villages, each beneficiary was provided with 1,000 FCFA (or about $2) per day for a maximum of 25,000 FCFA per month. Food payments consisted of commodities similar to those typically consumed in the region (see table 2 for a full list of food commodities provided in the ten countries). In this context, Hoddinott et al. (2014) assessed the relative impacts of cash versus food transfers through a randomized design using a single-difference evaluation approach (no baseline data was available). The results are differentiated by pre-harvest (or “hunger”) season in July and the post-harvest season in October.

The PSNP in Ethiopia is a widely studied flagship safety net program. Launched in 2005, the PSNP provides predictable transfers to about 7.5 million food insecure people for a period of 6 months each year. Transfers are delivered through two components: public works that provide temporary employment at a monthly wage of $16 (raised several times as a result of food prices), and an unconditional arm of direct support for households with limited working capacity. Sabates-Wheeler and Devereux (2012) compared the impacts of those cash and food interventions on food-gap and income using a single difference estimation strategy.

In Uganda, a program was implemented by the World Food Programme (WFP) in the fragile Karamoja region to support early childhood development. The program is a conditional food or cash transfer, where the former were provided as take-home rations, while the latter included a cash transfer of about $12 per child for each 6-week cycle (Gilligan and Roy 2013). Food transfers were distributed through the WFP’s general food distribution program, and cash transfers were added electronically to a card and retrieved from mobile money agent. Gilligan and Roy (2013) estimate the impacts of the program on child cognitive and non-cognitive development.
In Ecuador, a three-modality transfer program (featuring cash, food, and vouchers) was designed to address the food security and nutrition needs of Colombian refugees and to support their integration into Ecuadorian communities. The value of the monthly transfers was standardized across all treatment arms at $40 per month per household. The cash was transferred monthly onto preprogrammed ATM cards. Cash transfer households were able to retrieve the cash in bundles of $10. Vouchers were given in denominations of $20, redeemable for a list of nutritionally approved foods at identified supermarkets. Hidrobo et al. (2014) use the experimental design of this program to assess the impacts of transfers using an Analysis of covariance (ANCOVA) approach.

Aker (2013) evaluates the impacts of unconditional cash transfers and vouchers in the context of a complex emergency in the Democratic Republic of Congo. The program was part of the humanitarian response to internally displaced persons living in informal camps. Cash transfers and vouchers totaling $130 were provided over a seven-month period in three installments.19 Cash was directly deposited into an interest-free account at the office of a local cooperative located in the regional market center. Vouchers were used in fairs, including 122 vendors and 4 primary schools in the area. Impacts were estimated using a difference-in-difference with fixed effects.

In Yemen, an emergency intervention to address seasonal food shocks was carried out in 136-village clusters that were randomly assigned to receive either a food or a cash transfer. Over the course of seven months, households received 3 food baskets. Over the same time period, households in the cash treatment group received 3 cash transfers equal to the local value of the food basket (about $50). Cash transfers were distributed through ID cards from the Yemen Post and Postal Savings Corporation, and food transfers were delivered through the Ministry of Education. Schwab (2013) compared the impacts of food and cash transfers using single difference, difference-in-difference, and ANCOVA estimation strategies.

A pilot intervention was implemented in Cambodia as an extension to an existing food for education program. Food was provided through two modalities: early morning meals and take-home rations. The expansion included a cash scholarship to test the relative efficacy of the take-home rations and cash scholarships. Both transfers were conditional on a minimum attendance rate of 80 percent. The cash transfer was $60 per year and the take-home ration consisted of 10 kg of unfortified rice per month. Barker et al. (2014) estimated the comparative impacts using a difference-in-difference approach.

In post-tsunami Sri Lanka, an emergency cash transfer program was designed to support the needs of affected villages as part of the broader food-based operation run by WFP. The cash transfer pilot disbursed an average amount of US$2.44 per week. Food rations were specified for weekly amounts but in most cases transfers were given in bulk. Cash transfers were distributed by the Samurdhi Authority in
two-week installments through bank accounts. Sharma (2006) compared the relative impacts using a difference-in-difference estimation strategy.

In Bangladesh, Ahmed et al. (2010) examined the efficacy of four different programs utilizing food-only, cash-only, and combined cash and food transfers. For the purpose of this paper, we examine only the two “pure” cash and food transfer interventions, namely the food-based Income Generating Vulnerable Group Development (IGVGD) and the cash-based Rural Maintenance Program (RMP). The IGVGD program exclusively targeted poor women, who receive a monthly food ration over a period of 24 months. The RMP also targeted divorced, widowed, separated, or abandoned women with 4 years of employment to maintain rural roads. The RMP participants received a monthly net salary of $30. Impacts of both programs were estimated using propensity score matching techniques.

The specific features of the interventions are summarized in table 1, and describe the program types, modalities, transfer size, transfer as a percentage of household pre-program expenditures, frequency of distribution, household size, duration, delivery mechanisms, sample size, and evaluation method.

**Impacts**

The compilation of evaluation covers a number of dimensions. Table 2 lists 15 of them, including their availability across the examined countries. In most cases, cash interventions were conducted as a smaller-scale addition to preexisting large food-based programs, or as part of crisis responses. It is somewhat natural, therefore, that the interventions tend to be generally food security-oriented. Indeed, about half of the indicators are food and nutrition-related, while indicators on poverty, income, and assets were collected in a more limited number of cases.20

Importantly, there is a rich body of literature and debate on the comparative performance of indicators and measurement techniques for food security and nutrition, particularly as they relate to access and utilization of food (Dary and Imhoff-Kunsch 2012; Barrett 2010; Wisemann et al. 2009; Webb et al. 2006). The selected studies show that, in absolute terms and across basically all indicators, cash, food, and vouchers are effective in meeting program objectives. They also rejected the myths of food being resold on markets or cash being spent on non-desirable items.21 The following subsections examine the relative impacts of cash and food transfers on food consumption, availability of calories, dietary diversity, poverty, and malnutrition.

One of the most widely used indicators in the sample is food consumption. Adopted in 7 out of the 10 countries, the measure accounts for expenditures or value of food consumed at the household level. Figure 1 shows the difference in impacts on food consumption in Yemen, Cambodia, Mexico, Ecuador, Sri Lanka, and Bangladesh. Differences are expressed in percentage points (namely, average...
Table 1. Descriptive Statistics of Impact Evaluations

<table>
<thead>
<tr>
<th>Name</th>
<th>Country</th>
<th>Program type*</th>
<th>Modality</th>
<th>Cash size (US$)</th>
<th>Food basket</th>
<th>Size as % of Pre-program HH exp.</th>
<th>Frequency of Transfer</th>
<th>HH Size</th>
<th>Exposure</th>
<th>Years of data</th>
<th>Delivery mechanism</th>
<th>Sample Size (HHs)</th>
<th>Evaluation Methods</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAL</td>
<td>Mexico</td>
<td>CT, UT</td>
<td>Cash, Food</td>
<td>13</td>
<td>F&lt;sup&gt;1&lt;/sup&gt;</td>
<td>11.5</td>
<td>Monthly (cash), bi-monthly (food)</td>
<td>4.2</td>
<td>trial 1 year</td>
<td>2004–2005</td>
<td>Biometric debit cards</td>
<td>5,028</td>
<td>DD</td>
<td>Cunha (2014)</td>
</tr>
<tr>
<td>Zinder project</td>
<td>Niger</td>
<td>PW, UT</td>
<td>Cash, Food</td>
<td>50</td>
<td>F&lt;sup&gt;2&lt;/sup&gt;</td>
<td>11.5</td>
<td>Bi-weekly</td>
<td>7</td>
<td>6 months</td>
<td>July to October 2011</td>
<td>Mobile ATMs, smart cards</td>
<td>2,209</td>
<td>SD</td>
<td>Hoddinott et al., (2014)</td>
</tr>
<tr>
<td>PSNP</td>
<td>Ethiopia</td>
<td>PW, UT</td>
<td>Cash, Food</td>
<td>16.2</td>
<td>F&lt;sup&gt;3&lt;/sup&gt;</td>
<td>N/A</td>
<td>Monthly</td>
<td>5</td>
<td>6 months per year</td>
<td>2006–2008</td>
<td>N/A</td>
<td>960</td>
<td>SD</td>
<td>Sabates-Wheeler and Devereux (2012)</td>
</tr>
<tr>
<td>Early Childhood Development</td>
<td>Uganda</td>
<td>CT</td>
<td>Cash, Food</td>
<td>12</td>
<td>F&lt;sup&gt;4&lt;/sup&gt;</td>
<td>N/A</td>
<td>6-week cycle</td>
<td>6.22</td>
<td>12 months</td>
<td>October 2010 to April 2012</td>
<td>Mobile money cards</td>
<td>2,450 (est.)</td>
<td>ANCOVA</td>
<td>Gilligan and Roy (2013)</td>
</tr>
<tr>
<td>Colombian refugees project</td>
<td>Ecuador</td>
<td>CT</td>
<td>Cash, Food, Vouchers</td>
<td>40</td>
<td>F&lt;sup&gt;5&lt;/sup&gt;</td>
<td>16</td>
<td>Monthly</td>
<td>3.8</td>
<td>6 months</td>
<td>April to November 2011</td>
<td>ATM card</td>
<td>2,122</td>
<td>ANCOVA</td>
<td>Hidrobo et al. (2014)</td>
</tr>
<tr>
<td>IDP's project</td>
<td>Democratic Republic of Congo</td>
<td>UT</td>
<td>Cash, Vouchers</td>
<td>18.57</td>
<td>V&lt;sup&gt;6&lt;/sup&gt;</td>
<td>18.96</td>
<td>Bi-monthly</td>
<td>5.5</td>
<td>7 months</td>
<td>September 2011 to March 2012</td>
<td>Bank accounts</td>
<td>252</td>
<td>Fixed effects</td>
<td>Aker (2013)</td>
</tr>
<tr>
<td>Unconditional safety net</td>
<td>Yemen</td>
<td>UT</td>
<td>Cash, Food</td>
<td>49</td>
<td>F&lt;sup&gt;7&lt;/sup&gt;</td>
<td>N/A</td>
<td>Bi-monthly</td>
<td>7.9</td>
<td>6 months</td>
<td>2011–2012</td>
<td>ID card via Postal Savings Corporation</td>
<td>3,353</td>
<td>SD, ANCOVA, DD, DDD</td>
<td>Schwab (2013)</td>
</tr>
<tr>
<td>Scholarship pilot program</td>
<td>Cambodia</td>
<td>CT</td>
<td>Cash, Food</td>
<td>5</td>
<td>F&lt;sup&gt;8&lt;/sup&gt;</td>
<td>2.5</td>
<td>Monthly</td>
<td>6</td>
<td>10 months</td>
<td>August 2011 to August 2012</td>
<td>On-site manual distribution</td>
<td>4,091</td>
<td>DD</td>
<td>Barker et al. (2014)</td>
</tr>
</tbody>
</table>

Continued
Table 1. Continued

<table>
<thead>
<tr>
<th>Name</th>
<th>Country</th>
<th>Program type</th>
<th>Modality</th>
<th>Cash size (US$)</th>
<th>Food basket</th>
<th>Frequency of Transfer</th>
<th>HH Size Exposure</th>
<th>Years of data</th>
<th>Delivery mechanism</th>
<th>Sample Size (HHs)</th>
<th>Evaluation Methods</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTPP</td>
<td>Sri Lanka</td>
<td>UT</td>
<td>Cash, Food</td>
<td>9.8</td>
<td>10</td>
<td>Bi-weekly (cash), bi-monthly (food)</td>
<td>3.8</td>
<td>3 months</td>
<td>Samurdhi Bank</td>
<td>1.357 s</td>
<td>DD</td>
<td>Sharma (2006)</td>
</tr>
<tr>
<td>IGVGD RMP</td>
<td>Bangladesh</td>
<td>UT, PW</td>
<td>Cash, Food</td>
<td>19.7</td>
<td>10</td>
<td>Bi-monthly (cash), monthly (food)</td>
<td>4.6</td>
<td>2–4 years</td>
<td>Public banks</td>
<td>1.200</td>
<td>PSM</td>
<td>Ahmed et al. (2010)</td>
</tr>
</tbody>
</table>

Note:

1Seven basic items—enriched corn flour, rice, beans, dried pasta soup, biscuits, fortified milk powder, and vegetable oil—and two to four supplementary items (including canned sardines, canned tuna fish, dried lentils, chocolate, breakfast cereal, or corn starch.

23.5 kg of grain (primarily maize in the first transfer period and sorghum in the second), 0.72 kg of pulses (cowpeas, red beans, or lentils), 0.14 kg of vegetable oil, and 0.035 kg of salt.

33 kg of cereals, plus pulses and oils.

4Food basket of approximately 1,200 calories, includes corn soy blend (“CSB” – highly fortified with iron among other nutrients), vitamin-A fortified oil, and sugar.

5The food basket was valued according to regional market prices at $40 and included rice (24 kilograms), vegetable oil (4 liters), lentils (8 kilograms), and canned sardines (8 cans of 0.425 kilograms) (voucher: The list of approved foods consists of cereals, tubers, fruits, vegetables, legumes, meats, fish, milk products, and eggs).

6three food fairs, where participants could get palm oil, sugar, cassava flour, beans, rice, vegetable oil, dried fish, salt, potatoes and peanuts.

7For an average household size of seven persons is 50 kg of wheat flour and 5.0 liters of vegetable oil.

810 kg of rice per month.

91.4 kg Rice, 1.4 kg Wheat flour, 0.42 kg Pulses, 0.14 kg Oil, 0.14 kg Sugar, 0.14 kg Corn soy blend.

10up to 20 kilograms (kg) of wheat or 16 kg of rice per month. *CT = conditional transfer, UT = unconditional transfer, PW = public works.
impacts among food-receiving households minus those among cash-recipients), with negative values indicating the cases for which cash is more effective, and vice versa.

Table 2. Indicators Included in the Evaluations

<table>
<thead>
<tr>
<th></th>
<th>Mexico</th>
<th>Niger</th>
<th>Ethiopia</th>
<th>Uganda</th>
<th>Ecuador</th>
<th>Congo</th>
<th>Yemen</th>
<th>Cambodia</th>
<th>Sri Lanka</th>
<th>Bangladesh</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total consumption</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Food consumption</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Non-food consumption</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Food gap</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Food diversity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Caloric intake</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Anthropometric measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Assets</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Poverty headcount</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Labor market</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>participation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sickness probability</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Anemia</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>School dropout rates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Cognitive development</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Figure 1. Relative Impacts of Food versus Cash Transfers on Food Expenditures (Difference in Percentage Points).
The figure shows that only in Ecuador were impacts of food consumption larger for food-receiving beneficiaries, including relative to both cash and voucher transfers. In Yemen, Cambodia, Mexico, Sri Lanka, and Bangladesh the impacts on food consumption are higher for cash than for food-beneficiary households. In three cases—Yemen, Cambodia, and Sri Lanka—the difference is double digit. For Mexico and Ecuador, the difference in impacts is, however, not statistically significant.²²

Measures for quantifying calorie intake may present additional information regarding the difference in impacts on per capita food availability. Figure 2 sets out the impacts of food and cash transfers for programs in Sri Lanka, Yemen, Mexico, Ecuador, and Bangladesh.

In contrast with food consumption measures, food transfers have a larger impact on calorie intake relative to cash in most contexts. In Ecuador, the larger effect on calories from food was mainly due to larger increases in cereal consumption (which represented 41 percent of households’ caloric intake). In Yemen, higher caloric consumption from food stemmed from the basket composition, including wheat and oil. In the case of Sri Lanka, the impact is negative but not significant. After disaggregating these impacts geographically, Sharma (2006) notes that the impacts only decline significantly for one sub-region. The reason for this spatial difference is that

Figure 2. Relative Impacts of Food versus Cash Transfers on Per Capita Calorie Intake (Difference in Percentage Points).

![Figure 2. Relative Impacts of Food versus Cash Transfers on Per Capita Calorie Intake (Difference in Percentage Points).](https://example.com/figure2.png)
the household baseline survey was conducted the week after Muslim and Hindu festivities. Additionally, the negative effect can be explained by a change in diets, that is, a shift in consumption from highly caloric foods to diets of higher quality (e.g., eggs, meat). In the case of Bangladesh, cash transfers had a larger impact on food consumption. One possible explanation is that the size of the cash transfer was 70 percent higher than the food transfer. Ahmed et al. (2010) address this difference by comparing the change in marginal propensity to consume food, which shows consistent results.

Indicators that measure dietary diversity can provide insightful information on the quality of consumption patterns, in addition to its quantity. Three indicators are used in the evaluations, namely the Dietary Diversity Index (DDI), Food Consumption Scores (FCS) and Household Dietary Diversity Score (HDDS). Difference for FCS is available for Cambodia, Niger (July and October), Ecuador, and Yemen. Results are mixed, with cash being more effective in three cases (Ecuador’s cash and vouchers arms and Yemen), and food in the other three (Cambodia and Niger, both seasons). In Ecuador, the larger increase in dietary diversity for vouchers was mainly due to larger increases in the number of days that participants consumed vegetables, eggs, milk, and dairy. Similar effects of transfers were noted for the DDI, which included the same sample of countries except Cambodia. One reason that the cash recipients had less diverse diets lies in their choice of using a significant proportion of their transfers to buy grains in bulk, the least expensive form of calories present on local markets. As pointed out by Hoddinott et al. (2014), such a purchasing strategy was a reflection of uncertainty regarding future food prices.

The two studies in Congo and Ecuador also allowed for comparing the HDDS index among cash and voucher-receiving arms. In Congo, vouchers had a large impact on households’ dietary diversity (by 3.36 food groups), a 15 percent increase. Despite the comparatively lower impact of cash transfers, the difference in impact between cash transfers and vouchers is not significant. In Ecuador, the impact among cash and voucher recipients is fairly similar.

The food gap measure indicates the decrease of months of food shortage. In the case of Ethiopia, a two-year exposure to food rations led to less months of food shortage compared to households participating in cash transfers (public works). In Uganda, among cash- and food-treated households there was a reduction of 0.6 and 0.4 months of food insecurity, respectively. However, the difference is not statistically significant.

Some countries present data on the relative impacts of transfers on short-term and long-term nutrition-related dimensions. In Mexico, both food and cash transfers increased the intake of micronutrient (iron) amongst children by 1.61 mg and 1.10 mg, respectively. However, the difference is not statistically significant. The same pattern holds for increases in zinc and vitamin C. Similarly, anemia
prevalence was reduced by 2 percent in food-receiving households and 4 percent in the cash-receiving households.

In Uganda, the prevalence of severe underweight was 3.8 percent lower in the cash group than in the food group, a result significant at the 10 percent confidence level. Cash transfers reduced anemia prevalence by about 10 percentage points for young children (also in this case at 10 percent confidence level). In this context, food transfers had no significant impact. Finally, in Cambodia neither treatment modality in the food-cash scholarship program had significant impacts on anthropometric indicators, possibly because of the small transfer size and short exposure to treatment.

In some contexts cash and food transfers had an income multiplier effect on beneficiary households. In Ethiopia, PSNP food beneficiaries had a positive and significant income growth of 59.7 percent (at a 1 percent confidence level). In Cambodia, treatment households (both food and cash recipients) increased their net disposable liquidity by about $54 over the course of the year. In Bangladesh, both food and cash transfers increased income significantly: the cash-based RMP increased beneficiary income by 31.4 percent and the food-based IGVGD by 27.8 percent.

Both food and cash transfers reduced poverty in Mexico and Bangladesh. Skoufias et al. (2008) showed the impact of cash and food transfers on the head-count ratio as measured by the food poverty line. Similarly, Ahmed et al. (2010) estimated the impacts of cash transfers from the RMP program and food transfers from the IGVGD on the extreme poverty headcount ratio.

In both instances, food transfers had larger impacts, with a difference on 3.8 and 1.94 percentage points in Bangladesh and Mexico, respectively. The overall impacts of transfers on the poverty gap are larger. In the context of Mexico, food transfers decreased the poverty gap by 22.3 percent and cash transfers by 18.9 percent; moreover, the severity of poverty decreased by 27.8 percent and 22.97 percent, respectively. Skoufias et al. (2008) argue that the PAL transfer—equivalent to 11.5 percent of pre-transfer level household consumption—may have generated a multiplier effect that led to a reduction by 13 to 15 percent in the headcount poverty rate in two years. In the case of labor market participation, both food and cash transfer had a negative impact in agricultural activities, but a positive one in non-agricultural activities. Cash-receiving households increased non-agricultural activities by 7.1 percent and food-receiving households by 5.8 percent.

Based on the results presented in this section, figure 3 summarizes impacts by transfers and indicators. On average, impacts tend to be balanced across modalities, with no systematic dominance of one transfer over the other.
Figure 3. Summary of Relative Impacts by Transfer Modality and Indicator.

Note: Bars in light gray refer to percentage points, dark gray bars refer to changes in specific indicator values. Level of significance is indicated by the asterisks (* at the 90 percent level, ** at the 95 percent level, *** at the 99 percent level).
 Costs

Most of the examined evaluations also conducted a cost analysis, including Mexico, Niger, Ecuador, Democratic Republic of Congo, Yemen, and Bangladesh. In Mexico, logistics costs of moving commodities from warehouses to villages have been estimated to be about 30 pesos per box; this is equivalent to nearly 20 percent of the wholesale cost of the transfer (150 Mexican pesos, or approximately USD 15). Since cash transfers were distributed in the same way as under Mexico’s Oportunidades program, based on such data Cuncha (2012) estimated that it costs 2.4 percent of the transfer amount to deliver cash to recipients. Therefore, food is about ten times more costly than cash transfers, although it is of higher value (indeed, the 150 pesos cash transfer could only purchase about 73 percent of the in-kind basket; see table 3).

In Niger, food was about three times more expensive to implement than cash payments. However, Hoddinott et al. (2014) excluded costs that were common to both the food and cash payments, such as expenses associated with implementing the public works, identifying the beneficiaries, program sensitization, identification of implementing partners and contract negotiations with partners selected to implement this intervention. Some cash-specific costs were not considered, like fixed costs associated with setting up the cash delivery system. For example, each smart card used for the cash payments cost USD 6.00, while there were additional costs associated with computer programs needed to dispense payments through mobile ATMs.

In the case of Ecuador, food costs were about four times higher than for cash and vouchers. The cost of physical materials associated with vouchers, such as printing, is trivial. However, significant staff costs are associated with supermarket selection, the negotiation of contracts with individual supermarkets, and voucher reconciliation and payment. These staff costs account for nearly 90 percent of the cost of implementing the voucher component of the intervention. The cost of generating the debit cards was the main cost item in the cash transfer. The food transfer was significantly more expensive because of the cost of transporting the food to the

<table>
<thead>
<tr>
<th>Country</th>
<th>Food</th>
<th>Cash</th>
<th>Voucher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico</td>
<td>$2.29</td>
<td>$0.31</td>
<td></td>
</tr>
<tr>
<td>Niger</td>
<td>$12.91</td>
<td>$4</td>
<td></td>
</tr>
<tr>
<td>Ecuador</td>
<td>$11.50</td>
<td>$3.03</td>
<td>$3.30</td>
</tr>
<tr>
<td>Dem. Rep. of Congo</td>
<td>$11.34</td>
<td>$14.35</td>
<td></td>
</tr>
<tr>
<td>Yemen</td>
<td>$10.37</td>
<td>$4.09</td>
<td></td>
</tr>
</tbody>
</table>

distribution sites and the rental of storage facilities. Taking bulk items and repackaging them for distribution is also very costly, accounting for approximately 30 percent of the cost of distributing the food ration (Hidrobo et al. 2014).

Interestingly, Margolies and Hoddinott (2014) noted that food logistics costs also hinge on the location of food distribution points—that is, higher costs can be in part explained by delivering food directly to beneficiaries’ communities, such as for security reasons in Yemen; yet cash beneficiaries in Yemen had to collect the transfer at post offices, which meant higher transaction costs. In Ecuador, food distribution sites were located farther than cash and voucher payment points, thus increasing private costs (time and money); in Uganda and Niger, there appears to be no difference in transaction costs since both transfers were distributed at the village-level. In general, there appears to be a trade-off between costs for the implementer and those for beneficiaries: as payment or distribution points get closer to beneficiaries, costs for the implementer get higher while the transaction costs for beneficiaries dwindle. In other words, programs that seem less expensive could be so because the cost of obtaining benefits had been shifted from the implementer to the beneficiary.

The studies also raise the question of how to measure the value of food, whether at local market prices or procurement costs, with the latter being a more realistic cost (including for the different commodities that compose the basket). For two cases, Ecuador and Yemen, Margolies and Hoddinott (2014) estimated the full cost of cash and food transfers, with costs for food taken at the procurement level. In Ecuador, the procurement costs for food were higher than their local market value: indeed, accounting for the local procurement of most canned fish, rice, lentils and oil, and including the international procurement of some oil and lentils, it turns out that it cost USD 46.76 to provide a transfer that is locally valued at USD 40. This led to a total cost of providing food of USD 58.25 (USD 46.76 plus 11.46), which even exacerbated the cost differences, that is, total cost for cash is USD 42.99 while for vouchers it is USD 43.27 (the value of both voucher and cash transfer is, by definition, USD 40). Indeed, the difference between food and cash is now USD 15.26 per transfer compared to USD 8.47 (i.e., USD 11.46 minus USD 2.99) when transfer values were excluded.

In Yemen, food was more than twice as costly as cash transfers. Yet the difference becomes significantly smaller if additional cost items are included. For instance, households faced significant transaction costs (travel) for receiving the cash transfer. By factoring beneficiary cost to collect transfers, this raises the per-transfer cost (excluding the value of the transfer) of cash to USD 8.22. For food transfers, the addition of beneficiary costs raises the per-transfer cost (excluding the transfer value) to $11.35 (Schwab 2013). Moreover, market conditions were such that it was possible to procure for USD 39 a food basket locally valued at USD 49. Considering those
factors, the cost difference between food and cash cost even reversed, with cash being USD 2.8 more expensive than food (figure 4).

In the Democratic Republic of Congo, vouchers were more expensive than cash transfers. The cost breakdown shows that staff time represents the largest percentage of costs for both interventions, followed by transport and voucher printing (for the voucher intervention) and account-opening fees (for the cash intervention). Yet since the account opening fees are a one-time, fixed cost, if the program were to continue cash transfers with existing beneficiaries, the cost per cash program recipient would have only been USD 6 or 8 less expensive per program recipient (Aker 2013).

In Bangladesh, Ahmed et al. (2010) considered the costs for delivering cash and food transfers. For cash, only the bank transaction cost was considered, thus generating a very low cost for cash transfers (0.00115 per taka delivered). In the case of food, the analysis considered procurement costs as well as costs incurred at ports, losses, internal transport, storage, and handling, for a total cost of 1.2 taka per 1 taka transferred.

Yet few studies have explicitly considered costs and impacts (or cost-effectiveness) of a program, hence quantifying the possible trade-offs between the two. As illustrated by a maternal-child health program in Honduras, “. . . it cost 1.03 lempiras to deliver 1 lempira of income transfer in the form of a cash-like coupon, while it cost 5.69 lempiras to deliver the same income transfer in the form of food. However,

**Figure 4.** Difference in Total Costs (USD) between Transfer Modalities, With and Without Procurement (Food Minus Cash Costs).
the cash transfer had no effect on child’s calorie consumption nor on use of the health centers, while the food transfer increased both.” (Rogers and Coates 2002). In Bangladesh, Ahmed et al. (2010) estimated that the cost required to increase 100 kcal to beneficiaries is roughly similar for cash and food transfers (USD 3.28 and USD 3.21, respectively), while the cost for increasing household income by 100 taka is lower for food (USD 0.68) than cash transfers (USD 1.27). In Ecuador it was estimated that food transfers cost twice as much as vouchers to boost calorie intake by 15 percent, and were 5 times as costly as a way of increasing dietary diversity and quality (Hidrobo et al. 2014). Table 4 details the results for Ecuador.

In sum, cash and vouchers seem generally more efficient than food transfers. However, conclusions should be drawn with caution. Methods and approaches vary greatly, hindering intra- and cross-country comparisons. Most of the studies assess, for instance, food costs at local market prices instead of procurement costs. In other words, the availability of tools for transparent, comparable, and systematic cost analysis is limited. The development of such tools could help provide a more accurate and realistic account of comparative program performances.

### Discussion

The evaluations reviewed in the previous section prompt two questions: first, are those studies helping to advance the debate meaningfully? The answer is yes, definitely so, including by providing credible evidence, fresh perspectives and gearing the discussion toward results-oriented perspectives. In line with Hoddinott (2013), “Policy and interventions design discussion need to shift from their current ideological and political focus to one that emphasizes greater precision in interventions objectives, more nuanced understanding of context, and greater emphasis on costing.” The studies not only provide solid evidence based on counterfactuals, but also offer a platform for discussing a number of related social protection design and policy matters.

However, are the reviewed evaluations conclusive about the debate? No, most probably not. Indeed, like any evaluation method, experiments have inherent

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Food</th>
<th>Cash</th>
<th>Voucher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption</td>
<td>$10.78</td>
<td>$3.79</td>
<td>$3.81</td>
</tr>
<tr>
<td>Calories</td>
<td>$10.78</td>
<td>$7.58</td>
<td>$4.50</td>
</tr>
<tr>
<td>Household dietary diversity score</td>
<td>$28.75</td>
<td>$11.36</td>
<td>$8.25</td>
</tr>
<tr>
<td>Dietary diversity index</td>
<td>$15.68</td>
<td>$3.25</td>
<td>$2.91</td>
</tr>
<tr>
<td>Food consumption score</td>
<td>$17.25</td>
<td>$4.13</td>
<td>$3.09</td>
</tr>
</tbody>
</table>

Source: Hidrobo et al. (2014).

Gentilini
advantages and limitations. While not all the studies were RCTs, findings from those methods are particularly relevant when interpreted in light of the very specific circumstances that define a given setting, although that may limit extrapolations and lessons for different scenarios (external validity).

Importantly, our review showed that context and design matter. For example, the impacts tend to be associated with factors such as the marginal propensity to consume food, the duration of the program, the frequency of transfers, and perceived risks (e.g., of prices in the context of Niger), all of which affect purchasing and consumption behaviors. The composition of food baskets also plays an important role, including having a direct bearing on household dietary diversity and calorie availability. Therefore, a key issue that emerges from the evaluations is the need to interpret transfer selection as part of an organic decision-making framework. In particular, the choice of an optimal safety net should be guided by the specific objectives pursued, the peculiar contexts where programs are implemented, the choice of key design parameters (e.g., targeting method, conditionality, transfer size, duration, etc.), and the selection of modalities (including based on technical issues such as the functioning of markets, etc.). Transfer performance is ultimately and largely a function of those factors, instead of an inherent superiority of one modality over the other.25

Of course, a range of other elements help shape the program environment, such as the role of theory (including how that shapes the views of decision-makers), the availability of local and international evidence, and political economy. Although conceptually desirable, it will probably be impossible to fully separate technical and evidence-related issues from those of political economy. As such, real-world decision-making should perhaps consider political economy as an integral feature of the debate, and not necessarily a mere “non-technical” issue. Also, it is interesting to note that while only a few years ago the use of technology seemed to give an edge to cash transfers, the adoption of technological solutions (e.g., on payments, monitoring, verification of identity, etc.) now also underpins a range of voucher and food-based programs, hence making the field more leveled and overall transparent.26

The paper also shows that results on effectiveness (impacts) seem considerably more robust than those on efficiency. Devising robust and standardized tools and methods for identifying, collecting, and analyzing cost data should be a key priority for the transfer debate. In this regard, it would be important for cost calculations to be based on a more nuanced understanding of supply chains and agricultural markets. Indeed, implementation models can vary considerably, depending on the specific approaches and actors involved at different points in the supply chains. In the case of school feeding, for instance, Gelli and Suwa (2014) noted that “. . . different approaches can even co-exist within the same country, where, for instance, programme implementation is owned by decentralised institutions (e.g. individual states in Brazil or India), or where agencies (. . .) are complementing the national
programmes (e.g. Ghana and Kenya), [or models] linking the provision of goods and services for school feeding to smallholder farmers and the community.” Those nuances need to be taken into account for credible cost analyses. Following Gelli et al. (2012), “... a holistic, system wide analysis of the goods, funds and information flows between the involved stakeholders and helps to identify the opportunities to achieve cost-effective and sustainable programs.” In a similar vein, the extensive logistics required for food-based programs—including procurement, transport, storage, and distribution—may posit particular risks for accountability, transparency, and “leakages”. In India, for example, it was estimated that diversion and corruption account for significant losses: in the early 2000s, about 58 percent of the food under the PDS program did not reach the intended beneficiaries (World Bank 2011). Such losses throughout the logistics chain should also be accounted for using comprehensive cost-effectiveness assessments of alternative modalities.

Because of program objectives and rationale, the evaluations included a range of food security-related measures; more comparative research around a wider set of dimensions, especially child malnutrition, would help advance research and practice. Value-based vouchers as a modality seem underexplored relative to their proven and possible performance, and so is the combinations of cash and food transfers. For instance, new evidence suggests considerable potential stemming from a mix of modalities, especially in protracted emergency contexts (Langendorf et al. 2014; Ruel and Alderman 2013). At the same time, attention should be paid to how objectives and metrics are set. For example, if a program is designed to pursue food security objectives, studies often do not capture what might be the side-effects on housing, clothing, debt repayment, etc. In other words, “choosing what is important” in terms of a study’s focus may be subjective, hence leaving some other important dimensions unexplored.

Conclusions

This paper reviewed key issues and evidence in the cash versus food debate, including a new generation of comparative evaluations undertaken in ten developing countries. Those studies represent a refreshing evidence-based approach that help inform and navigate a longstanding quandary.

The reviewed evidence shows that, in absolute terms, both modalities work. When compared to control groups, cash and food transfers (and vouchers when considered) bolstered improvements in a range of indicators such as food consumption, income, dietary diversity, poverty, and malnutrition. The paper also argues that, in relative terms, transfer modalities can lead to varied and mixed impacts over a range of dimensions. We observe a mild tendency of cash transfers to be more effective than food in enhancing food consumption (in five studies out of seven),
while food seems to outperform cash in increasing household caloric intake (in four evaluations out of six). However, overall effectiveness cannot be generalized and it depends not only on particular objectives, but also on the specific indicators used to measure those objectives.

Differences in design and context explain part of the difference in performance between alternative transfer modalities. Indeed, the impact of cash and food transfers can also differ in programs with identical design and contexts. Therefore, since transfer modalities can alter the impact pathways of an intervention, transfer selection should be considered as a key choice in safety net decision-making, similar to other program design features such as conditionality and targeting methods.

Costs will be an increasingly key factor in gauging performance, hence calling for more rigorous and systematic cost-effectiveness analyses than many of those currently available. In general, the reviewed studies show that cash transfers and vouchers tend to be more efficient than food-based interventions. Yet results should be interpreted with caution. Approaches for cost calculations are often not clearly described, rely on simplistic assumptions, and tend to be discretionary in the type of cost items considered. More standardized and robust approaches are required so that efficiency analyses match the higher standards of effectiveness offered by the examined impact evaluations.

Notes

Ugo Gentilini is a Senior Economist with the World Bank’s Social Protection & Labor Global Practice. He can be contacted at Ugentilini@worldbank.org.

1. See, for example, Andrew et al. (2014), Blattman and Niehaus (2014), Ozler (2013), Davis et al. (2012), and Independent Evaluation Group (2011).

2. A notable exception is the United States’ Supplemental Nutrition Assistance Program (SNAP), or ex-Food Stamp Program. As mentioned in the next section, SNAP has been subject to ample empirical scrutiny in comparison to cash transfers.

3. In most cases, commodities are provided by private or commercial partners. Yet retailers could also be public or a mix of public-private actors, such as the case of fixed-price outlets in Syria (Omamo et al. 2010).

4. In practice, commodity-based vouchers can also provide some room for choice. There are instances where beneficiaries can choose between commodities so long as their weight does not exceed the limit established by the program (Levan Tchatchua, personal communication).

5. Initiatives such as WFP’s Purchase for Progress program deliberately procure food from small farmers, hence spurring local economic gains among well-identifiable actors similar to vouchers’.

6. The program is the world’s largest food voucher scheme. According to USDA data, about 1 every 7 Americans, or 47.4 million individuals, benefited from the SNAP voucher program from 2013 to 2014. See: http://www.fns.usda.gov/pd/34SNAPmonthly.htm.

7. During the same period, the debate was also examined in the context of famine relief as laid out in the influential analysis by Dreze and Sen (1989, p.97).

8. An in-kind transfer is defined as extramarginal (or inframarginal) when it is larger (smaller) than what households would have normally consumed in the absence of the transfer. When extramarginal, the theory predicts that, under certain conditions, in-kind transfers would be more effective in augmenting food consumption than an equal cash transfer.
9. Explanations include, among others, a “labeling” effect inducing a sense of moral obligation to use in-kind transfers for their intended food consumption purpose (Senauer and Young 1986); the role of gender and decision-making behaviors in multi-adult households (Breunig and Dasgupta 2005); or alterations in household budgeting and planning of monthly purchases induced by in-kind transfers (Wilde and Ranney 1996).

10. For instance, since 2000 locally-procured food in developing countries soared by 45 percent, while untargeted bilateral food aid (i.e., food aid sold on local markets) now represents only about 3 percent of total flows. The entire volume of international food aid itself amounts to less than 0.25 percent of total global food production (Barrett et al. 2012). Gradually, innovations such as the use of smart cards, mobile phones, biometric devices, and e-monitoring tools were also adopted in food and voucher programs (Omamo et al. 2010).

11. The political economy and distributional aspects of the debate were indirectly captured by James Tobin in his 1970 Henry Simons Lecture at the University of Chicago. Indeed, while Tobin proposed large-scale means tested cash transfers to reduce poverty without interfering with the market determination of relative prices, he also wished to pair these with “non-market egalitarian distributions of commodities essential to life and citizenship” such as food stamps and other measures.


13. For an interesting view on how recent high food prices may have ultimately benefited the poor, see Headey (2014).

14. Counter to this logic, in some cases markets are (physically) brought to locations where a cash program is implemented. This is the case of fairs, for example, including areas where traders would not operate in the absence of the program. In such cases, cash-based transfers would be provided independently of the presence of working markets. The example of Congo discussed later in the paper falls under this category (Aker 2013).

15. Note that the issue of expressing and capturing preferences is notoriously difficult to handle, including a number of factors that may distort feedback such as who conducts the survey (e.g., founder of the program), expectations by beneficiaries, etc.

16. See some interesting initial guidance laid out in Ryckembusch et al. (2013), White et al. (2013), and Gelli et al. (2012).

17. Mexico benefited from two distinct evaluations by Cunha (2014) and Skoufias et al. (2008).

18. A new RCT is currently underway in Bangladesh, including compositions of cash, food, and nutrition training (Ahmed et al. 2013). Total sample size includes 5,000 households (4,000 of which are treated). Final results are expected to be released in late 2014.

19. Cash payments were made in three disbursements: September 2011 (US$90); November 2011 (US$20) and February 2012 (US$20). Program recipients had to travel to this market center to receive the cash transfer. Vouchers were provided in three installments, the first of which were commodity-based, while the last two were value-based.

20. In the case of Niger, Hoddinott et al. (2014) do not report on labour market participation and assets, although their survey includes that information.

21. In Mexico, more than half of the cash transfer was spent on food, out of which a quarter was devoted to nutritious food items such as fruits and vegetables. In Niger, cash recipients reported spending 70 percent of the transfer on food items, 10 percent on nonfood items, 9 percent on transfers to the households, saving 7 percent, and 3 percent to pay back loans. In Ecuador, cash beneficiaries used 83 percent of the transfers for food expenditures. The remainder was spent on nonfood expenditures (63 percent), shared with others outside the household (24 percent), and saved for later use (83 percent). In Congo, cash households used their transfers to purchase over six different categories of goods, health expenses, school fees, and debt reimbursement. In Yemen, cash households report spending 88 percent of their transfer towards repaying debts (5 percent), transportation (2 percent), and near zero on qat (14 out of 10,500 YER). In Uganda, the average cash beneficiary spent...
53 percent of the transfer on food (41 percent on stables), while 23 percent was allocated for nonfood goods and 16 percent of the cash was saved.

22. Note that for Mexico, estimates are based on Skoufias et al. (2008) examining the same program and data as Cuncha (2012). Skoufias et al. show the consumption indicator in percentages, while Cunha uses currency.

23. The Dietary Diversity Index reflects the number of different foods or food groups consumed over a given reference period. The Food Consumption Scores index measures the frequency of consumption of 8 food groups during the seven days before the survey. The Household Dietary Diversity Score indicates the number of food groups consumed in the previous seven days from 12 groups. It differs from the DDI in that frequency is measured across standardized food groups instead of individual food items.

24. Respondents reported a number of additional costs. For example, they indicated that on average it cost 480 FCFA (roughly 1 USD) to transport 100kg of cereals from the market to home, or otherwise 1,920 FCFA for the transfer period (four trips). This figure, however, does not take into account households' pooling transport costs, which could significantly reduce the per-household cost. The average cost for obtaining the food transfers by beneficiaries, as reported in household surveys, was only 60 CFA per trip.

25. In terms of comparability, it should be noted that any transfer comparison requires some basic level of functionality in implementation conditions (e.g., in markets, delivery systems, etc.). This somewhat obscures an important advantage of food transfers, which centers on its ability to operate in more extreme circumstances (e.g., absence of markets, etc.).

26. For example, WFP is currently supporting some 880,000 Syrian refugees in Lebanon through the use of electronic vouchers implemented in partnership with Mastercard. In Chhattisgarh, India, the PDS program distributes food by using portable technologies available among participating retailers and allowing people to buy commodities in their store of choice.

References


The Price Is Not Always Right: On the Impacts of Commodity Prices on Households (and Countries)

Daniel Lederman¹ and Guido Porto²

This paper provides an overview of the impact that one-time changes in commodity and other prices have on household welfare. It begins with a collection of stylized facts related to commodities based on household survey data from Latin America and Africa. The data uncovers strong commodity dependence on both regions: households typically allocate a large fraction of their budget to commodities, and they often also depend on commodities to earn their income. This income and expenditure dependency suggests sizable impacts and adjustments following commodity price shocks. The article explores these effects with a review of the relevant literature. The authors study consumption and income responses, labor market responses, and spillovers across sectors. The paper provides evidence on the relative magnitudes of various mechanisms through which commodity prices affect household (and national) welfare in developing economies. Commodity price changes, Poverty and welfare impacts, Net consumers and net producers. JEL codes: I30, O13, Q17

Energy prices rose more than 140% between 2000 and 2013, according to the World Bank’s energy-price index, while the food-price index rose by more than 80%. These increases in commodity prices motivated researchers to assess the implications for households and workers in developing countries, as well as policy responses. An example of such research is the World Bank’s $4 billion contribution to a joint fund with J.P. Morgan, launched on June 21, 2011, to help developing countries invest in commodity-price hedging instruments. Yet commodity prices have stalled since 2012, and observers are once again concerned about the social and economic consequences of declining commodity prices. Part of the concern is about the volatility of commodity prices, as much as the medium-term trends, per se.
Whether it is the trends or the volatility of commodity prices, knowledge about the effects of commodity-price fluctuations on household welfare is useful for policy as much as it is useful for understanding economic behavior. It is also important to have a sense of the magnitudes of the impacts involved.

The objective of this article is to put forth an economic framework for understanding how people in developing countries are affected by relatively permanent changes in commodity and other prices. The approach is to review the literature framed by a model that highlights issues relevant in the discussion of prices and welfare. The literature review canvasses research from various branches, including shocks to exchange rates, financial crises, and trade reforms. In all these cases, the shocks manifest themselves as price changes, and thus the findings are relevant for the discussion of commodity prices. We complement the literature review with a detailed analysis of household-survey data on commodity dependence in Latin America and Africa.

Since we are interested in the impact of commodity prices on household welfare and behavior, our framework builds on agricultural household models. This framework analyzes how households choose consumption bundles and how they earn their living. In some cases, especially in rural areas, income is generated by agricultural production (profits). In other cases, where the economy is more industrialized, wages and employment are the predominant sources of income, which can be affected by commodity prices. Since prices can affect both consumption and income, the simple framework opens the door to a discussion of how household welfare is affected by changes in commodity prices, both in the short- and long-run. The literature provides numerous empirical findings that can bring together various pieces of the puzzle illustrated by the framework.

The paper is structured as follows. First, we review the literature on the first-order impacts of price changes. These are the impacts that would take place if households adjust neither the quantities of commodities that they consume nor change the sources of income they receive in response to the price change. The article derives the commonplace net-consumer-versus-net-producer result, which establishes that a price increase hurts net-consumers and benefits net-producers.

Second, we discuss spillovers from commodity prices to other sectors. This refers to, for instance, whether households adjust all consumption choices in response to higher food prices, or whether only food expenditures are affected. We include a discussion of scenarios where commodities are used as inputs into the production of other goods. Consequently, changes in commodity prices (such as oil) may affect the cost, and thus the price, of certain goods (such as transport). Finally, we also report evidence on spillovers in employment and wages: changes in food prices can affect wages and employment not only in the food sectors (primary, agro-manufacturers) but also in non-food sectors.

Third, the article reviews how prices are transmitted to the local economy. Because commodities are traded in international markets, most of the shocks
associated with commodity prices are global. Consequently, a relevant question is to what extent do changes in international prices pass through to domestic markets?

Fourth, the article explores available evidence on consumption and production responses. Here, we review how consumers adjust the quantity and quality of the consumption basket, as well as other behavioral responses associated with consumption. We also explore some of the patterns of adjustment in production, particularly supply responses in agriculture.

It is important to mention at the outset that there are at least five related literature surveys. These include Winters, McCulloch, and McKay (2004) and Goldberg and Pavcnik (2004), both of which cover trade liberalization and poverty; Goldberg and Pavcnik (2007) and Harrison, McLaren, and McMillan (2011), which cover trade and the distribution of income; and Lederman (2013), which covers trade and inclusive growth. Our review complements this work by merging issues of poverty with issues of inequality. Furthermore, we have a somewhat narrower focus, as we are mainly interested in commodity prices, but we take a somewhat deeper look at some of these issues, all in the context of a general framework.

The rest of this paper is organized as follows. The following section presents an overview of household commodity dependence in Latin America and Africa. Using various household income and expenditure surveys from Latin American and African countries, we estimate the average budget shares spent on commodity-related sectors (mostly food and energy), and the average income share derived from commodity-related sectors (agriculture, forestry, fishing, mining, food manufactures, and energy). The data reveal strong commodity dependence on both regions. The next section introduces the analytical framework and reviews the literature that measures those effects using data from developing countries. The implications of price changes for net-consumers and net-producers are central to this analysis, with labor markets, various types of spillovers across markets, and imperfect price transmission of border prices all playing fundamental supporting roles in the story. The subsequent section turns to departures from first-order effects and studies the role of household adjustments in production and consumption. This is followed by a simple example of the magnitudes of the different channels identified in the paper using Mexican data, while the final section concludes.

Commodity Dependence: Stylized Facts

This section characterizes the dependence on commodities of households in developing countries. To study dependence on consumption, we focus on expenditure shares spent on items related to commodities, mainly food and energy. Households’ dependence on commodities on the income side is then measured by income shares in agriculture (including forestry and fishing), mining, food manufactures (beverages,
food, and tobacco), and energy. An important aspect of the analysis is related to poverty and inequality, because dependence on commodities either on the consumption or the income side can vary across quintiles of the income distribution during a given point in time.

Regarding commodity-dependence on the consumption side, table 1 presents the average expenditure shares on commodities for eight Latin American countries and 19 African countries. In Latin America, there is significant cross-country variation in the average share spent on commodities, ranging from an average share of 71.3 percent in Bolivia to 31.7 percent in Mexico. A common feature in all these countries is that the share of commodity expenditures decreases across quintiles. This is

<table>
<thead>
<tr>
<th>Quintiles</th>
<th>Argentina</th>
<th>Bolivia</th>
<th>Colombia</th>
<th>El Salvador</th>
<th>Mexico</th>
<th>Nicaragua</th>
<th>Panama</th>
<th>Peru</th>
<th>Burundi</th>
<th>Benin</th>
</tr>
</thead>
<tbody>
<tr>
<td>p/c expenditures</td>
<td>61.8</td>
<td>48.9</td>
<td>42.3</td>
<td>37.5</td>
<td>29.0</td>
<td>43.9</td>
<td>82.9</td>
<td>78.7</td>
<td>73.0</td>
<td>67.6</td>
</tr>
</tbody>
</table>

**Table 1. Commodity Expenditure Shares, Latin America and Africa**

<table>
<thead>
<tr>
<th>Quintiles</th>
<th>Argentina</th>
<th>Bolivia</th>
<th>Colombia</th>
<th>El Salvador</th>
<th>Mexico</th>
<th>Nicaragua</th>
<th>Panama</th>
<th>Peru</th>
<th>Burundi</th>
<th>Benin</th>
</tr>
</thead>
<tbody>
<tr>
<td>p/c expenditures</td>
<td>61.8</td>
<td>48.9</td>
<td>42.3</td>
<td>37.5</td>
<td>29.0</td>
<td>43.9</td>
<td>82.9</td>
<td>78.7</td>
<td>73.0</td>
<td>67.6</td>
</tr>
</tbody>
</table>

**Note:** Authors’ own elaboration based on household surveys.
expected due to Engel’s Law: richer households typically spend a smaller share of their total expenditures on food. In Bolivia, for example, the average commodity share is 82.9 percent in the first quintile and 54.4 percent at the top quintile. In Mexico, the average share of the bottom quintile is 40.6 percent and it is 20.8 percent at the top. In Colombia, the poorest households spend, on average, 80.1 percent on commodities; the richest households spend 41.6 percent. Differences in the populations covered by the surveys across countries (i.e., urban versus rural populations) can clearly account for part of the differences in the average commodity expenditure shares. Nevertheless, the data reveal a very high commodity dependence on the consumption side, especially for the poorest households. In Africa, there is also significant variation across countries, from a minimum average share of 42.7 percent in South Africa, to 84.4 percent in Tanzania, and a maximum of 89.9 percent in Madagascar. The commodity dependence on consumption seems higher in Africa than in Latin America, however. In most countries, commodity dependence in consumption is more prevalent at the bottom of the quintile distribution. This means that Africa and Latin America are both highly dependent on commodities and, moreover, this dependence declines with household income. This pattern explains the literature’s concerns about the poverty impacts of commodity price hikes.

Table 2 turns to dependence on the income side. The aim is to measure the share of household income derived from commodities, including income from employment (wages) and self-employment (for instance, agricultural home production). Starting with Latin American data for 2004, Table 2 reports the total combined share of income derived from commodities, which varies widely across countries. This comparison, however, is unlikely to be very useful because of differences in coverage. An example of this is Argentina, which despite being a country with a relatively important agricultural sector, the household survey suggests that only 3.2 percent of household income is, on average, derived from commodities. This is clearly more a manifestation of the urban focus of the survey than of the economy. In the discussion that follows, it is thus convenient to focus more on within-country comparisons across quintiles. A regularity in most of our Latin American data is that lower quintiles exhibit higher shares of income derived from “commodities.” The differences are sometimes important. In Paraguay, for instance, while the bottom quintile earns 44.4 percent of income from commodities, the top quintile earns only 14.7 percent.

We uncover similar patterns in the African data. The income commodity dependence in Africa is very high (higher than in Latin America) and varies across countries. The lower quintiles also tend to show higher commodity dependence. In Ghana, for instance, the share of income derived from commodities is 71.4 in the first quintile, and 27.1 in the top quintile. Table 2 shows that in the data, most of the dependence from commodities in Africa is related to dependence on agriculture.
Since the data from both regions indicates that poor households are on average highly dependent on commodities on both the consumption and income sides, the distinction between net-consumers and net-producers will become particularly

<table>
<thead>
<tr>
<th>Country</th>
<th>Quintiles p/c income</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Argentina</td>
<td>3.1</td>
<td>3.1</td>
</tr>
<tr>
<td>Bolivia</td>
<td>38.2</td>
<td>19.6</td>
</tr>
<tr>
<td>Brazil</td>
<td>22.3</td>
<td>14.8</td>
</tr>
<tr>
<td>Chile</td>
<td>17.7</td>
<td>15.6</td>
</tr>
<tr>
<td>Colombia</td>
<td>13.7</td>
<td>14.7</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>26.0</td>
<td>23.9</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>22.0</td>
<td>16.1</td>
</tr>
<tr>
<td>Ecuador</td>
<td>44.1</td>
<td>33.9</td>
</tr>
<tr>
<td>El Salvador</td>
<td>19.6</td>
<td>17.1</td>
</tr>
<tr>
<td>Guatemala</td>
<td>25.2</td>
<td>23.0</td>
</tr>
<tr>
<td>Honduras</td>
<td>43.7</td>
<td>34.8</td>
</tr>
<tr>
<td>Mexico</td>
<td>24.1</td>
<td>14.0</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>36.9</td>
<td>29.6</td>
</tr>
<tr>
<td>Panama</td>
<td>32.9</td>
<td>17.3</td>
</tr>
<tr>
<td>Paraguay</td>
<td>38.4</td>
<td>27.9</td>
</tr>
<tr>
<td>Peru</td>
<td>37.6</td>
<td>27.1</td>
</tr>
<tr>
<td>Uruguay</td>
<td>9.6</td>
<td>6.9</td>
</tr>
<tr>
<td>Venezuela</td>
<td>23.1</td>
<td>19.3</td>
</tr>
<tr>
<td>Burundi</td>
<td>54.0</td>
<td>57.9</td>
</tr>
<tr>
<td>Benin</td>
<td>60.6</td>
<td>58.3</td>
</tr>
<tr>
<td>Burkina Faso</td>
<td>73.9</td>
<td>69.4</td>
</tr>
<tr>
<td>Côte d’Ivoire</td>
<td>57.1</td>
<td>56.5</td>
</tr>
<tr>
<td>Cameroon</td>
<td>68.4</td>
<td>66.8</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>82.4</td>
<td>78.0</td>
</tr>
<tr>
<td>Ghana</td>
<td>67.5</td>
<td>65.2</td>
</tr>
<tr>
<td>Gambia</td>
<td>67.3</td>
<td>70.8</td>
</tr>
<tr>
<td>Guinea Bissau</td>
<td>68.4</td>
<td>72.6</td>
</tr>
<tr>
<td>Kenya</td>
<td>83.0</td>
<td>80.9</td>
</tr>
<tr>
<td>Madagascar</td>
<td>85.7</td>
<td>87.5</td>
</tr>
<tr>
<td>Mali</td>
<td>76.8</td>
<td>71.0</td>
</tr>
<tr>
<td>Malawi</td>
<td>58.8</td>
<td>61.0</td>
</tr>
<tr>
<td>Nigeria</td>
<td>88.5</td>
<td>75.3</td>
</tr>
<tr>
<td>Rwanda</td>
<td>77.7</td>
<td>67.3</td>
</tr>
<tr>
<td>Senegal</td>
<td>61.6</td>
<td>61.0</td>
</tr>
<tr>
<td>Tanzania</td>
<td>90.0</td>
<td>88.4</td>
</tr>
<tr>
<td>Uganda</td>
<td>71.2</td>
<td>73.0</td>
</tr>
<tr>
<td>South Africa</td>
<td>58.8</td>
<td>51.6</td>
</tr>
</tbody>
</table>

**Note:** Authors’ own elaboration based on household surveys.
relevant for the literature that attempts to assess the distributional and poverty impacts of commodity-price fluctuations.

As a preliminary step in this direction, it is worth analyzing the evolution of consumption and income dependence on commodities during a period of time when global commodity prices rose. The purpose of the empirical exercise is to describe the changes in income shares observed in the data from 2004 to 2008. Transitions in and out of dependency can be studied with panel data, which is unavailable in our sample countries. We can, however, build pseudo-panels of households using repeated cross-sections with the method proposed by Dang, Lanjouw, Luoto, and McKenzie (2011). Since only the income surveys in Latin America are repeated cross-sections, we focus here solely on income commodity dependence for Latin America. Due to sampling and measurement issues, one can obtain a distorted view of over-time trends by looking just at the household expenditures and incomes between cross sections. What is needed is a methodology that removes the effect of changes in household samples on the average household consumption and income shares derived from commodities.

To build the 2004–2008 pseudo-panel, we focus on households observed in 2004 and estimate commodity income shares for these households in 2008. This procedure requires the estimation of a regression model for commodity income shares separately for the 2004 and 2008 samples. We then use the estimates from the 2008 sample, together with the residuals from the 2004 regression to produce counter-factual panels. Intuitively, we predict average income shares using the observed covariates in 2004 and the estimated parameters for 2008.7

The results are reported in table 3. The first row reproduces the observed income shares in 2004 that were reported in table 2, while the second row shows the predicted 2008 shares. A key result is that income dependence on commodities increased in Chile, El Salvador, Honduras, Peru, and Uruguay, which is expected given the rise in global commodity prices. However, dependency remained roughly constant in (urban) Argentina and in the Dominican Republic; Brazil and Mexico experienced a slight decline in commodity income dependence, while in Costa Rica, Ecuador, Panama, and Paraguay, the decline was large. This international heterogeneity might be surprising given the commodity price boom, but differences are expected due to various factors discussed below.

It is also noteworthy that the dynamics of income dependence on commodities varies across quintiles within countries. In El Salvador, Honduras, and Uruguay, income dependence increased not only on average, but also for all quintiles. In Chile average dependence rose, but it decreased for lower quintiles and increased for higher quintiles. In contrast, in Peru, average dependence also rose, but it increased for lower quintiles and decreased for higher quintiles. In Argentina and the Dominican Republic, average dependence remained constant, but increased for low quintiles and declined for high quintiles in Argentina; it declined for low quintiles
and increased for high quintiles in the Dominican Republic. In Brazil, Ecuador, Mexico, and Paraguay, average dependence rose, declined for low quintiles, but increased for higher quintiles. Finally, in Costa Rica and Panama, dependence declined for the average as well as for all quintiles.

These results uncover several observations that are important for the subsequent analysis. First, the data show that households adjust after price shocks. In fact, the commodity income shares changed between 2004 and 2008, when global prices rose to record levels. Second, these adjustments were heterogeneous, not only across countries but also across quintiles of the distribution of income within countries. In what follows, the literature review aims to shed light on the impacts of

Table 3. Commodity Income Shares Methodology Using Repeated Cross-sections, Latin America, 2004–2008

<table>
<thead>
<tr>
<th>Country</th>
<th>Year</th>
<th>Quintiles p/c income</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>2004</td>
<td>3.1</td>
<td>3.1</td>
<td>3.2</td>
<td>3.2</td>
<td>3.5</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>4.2</td>
<td>3.8</td>
<td>3.2</td>
<td>3.0</td>
<td>2.5</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>2004</td>
<td>22.3</td>
<td>14.8</td>
<td>10.1</td>
<td>7.8</td>
<td>5.5</td>
<td>12.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>16.9</td>
<td>14.6</td>
<td>12.0</td>
<td>9.6</td>
<td>5.9</td>
<td>11.8</td>
<td></td>
</tr>
<tr>
<td>Chile</td>
<td>2003</td>
<td>17.7</td>
<td>15.6</td>
<td>13.4</td>
<td>10.8</td>
<td>9.0</td>
<td>13.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>16.0</td>
<td>16.0</td>
<td>14.0</td>
<td>12.6</td>
<td>10.9</td>
<td>13.9</td>
<td></td>
</tr>
<tr>
<td>Costa Rica</td>
<td>2004</td>
<td>26.0</td>
<td>23.9</td>
<td>19.1</td>
<td>13.1</td>
<td>8.9</td>
<td>18.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>18.9</td>
<td>20.3</td>
<td>17.5</td>
<td>13.6</td>
<td>8.1</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>2004</td>
<td>22.0</td>
<td>16.1</td>
<td>12.7</td>
<td>12.4</td>
<td>7.4</td>
<td>14.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>17.2</td>
<td>13.9</td>
<td>11.4</td>
<td>9.5</td>
<td>6.7</td>
<td>11.7</td>
<td></td>
</tr>
<tr>
<td>Ecuador</td>
<td>2004</td>
<td>44.1</td>
<td>33.9</td>
<td>24.5</td>
<td>17.5</td>
<td>11.5</td>
<td>26.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>32.3</td>
<td>28.6</td>
<td>24.4</td>
<td>19.8</td>
<td>12.3</td>
<td>23.5</td>
<td></td>
</tr>
<tr>
<td>El Salvador</td>
<td>2004</td>
<td>19.6</td>
<td>17.1</td>
<td>10.9</td>
<td>7.8</td>
<td>5.4</td>
<td>12.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>28.3</td>
<td>22.4</td>
<td>17.9</td>
<td>14.0</td>
<td>8.0</td>
<td>18.1</td>
<td></td>
</tr>
<tr>
<td>Honduras</td>
<td>2004</td>
<td>43.7</td>
<td>34.8</td>
<td>21.8</td>
<td>14.1</td>
<td>11.1</td>
<td>25.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>44.4</td>
<td>39.4</td>
<td>29.4</td>
<td>20.6</td>
<td>12.6</td>
<td>29.3</td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td>2004</td>
<td>24.1</td>
<td>14.0</td>
<td>11.0</td>
<td>7.1</td>
<td>6.4</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>18.9</td>
<td>13.9</td>
<td>11.1</td>
<td>9.4</td>
<td>7.0</td>
<td>12.1</td>
<td></td>
</tr>
<tr>
<td>Panama</td>
<td>2004</td>
<td>32.9</td>
<td>17.3</td>
<td>11.6</td>
<td>9.7</td>
<td>4.5</td>
<td>15.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>20.7</td>
<td>14.6</td>
<td>10.9</td>
<td>9.4</td>
<td>4.2</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>Paraguay</td>
<td>2004</td>
<td>38.4</td>
<td>27.9</td>
<td>22.5</td>
<td>17.1</td>
<td>15.5</td>
<td>24.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>33.0</td>
<td>26.9</td>
<td>22.9</td>
<td>19.1</td>
<td>14.7</td>
<td>23.3</td>
<td></td>
</tr>
<tr>
<td>Peru</td>
<td>2004</td>
<td>37.6</td>
<td>27.1</td>
<td>17.3</td>
<td>11.2</td>
<td>9.5</td>
<td>20.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>39.1</td>
<td>29.4</td>
<td>20.7</td>
<td>13.1</td>
<td>7.7</td>
<td>22.0</td>
<td></td>
</tr>
<tr>
<td>Uruguay</td>
<td>2004</td>
<td>9.6</td>
<td>6.9</td>
<td>4.7</td>
<td>3.8</td>
<td>3.5</td>
<td>5.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>14.9</td>
<td>14.5</td>
<td>12.8</td>
<td>11.3</td>
<td>8.6</td>
<td>12.4</td>
<td></td>
</tr>
</tbody>
</table>

Note: Authors’ own elaboration based on household surveys.
commodity price changes on household welfare and on household responses and adjustments in response to price signals.

Price Changes and Household Welfare

This section explores the impacts of price changes on household welfare. Our analytical framework builds on standard agricultural household models, as in Singh, Squire, and Strauss (1986). Although this framework has influenced a large body of research spanning various fields of economics, the ongoing discussion will modify it as needed to illustrate key issues associated with the welfare impacts of commodity price changes.

The framework analyzes the consequences of price variations from the viewpoint of households. Analyses of welfare changes resulting from such price signals tend to follow the approach pioneered by Dixit and Normal (1980), in which a household’s budget constraint is (at least initially) assumed to be fixed. Simply put, in equilibrium, household expenditures (including savings) have to be financed with household income (including transfers).

In turn, total household expenditures are, by definition, the sum of expenditures across goods, evaluated at their corresponding prices and quantities (for a given level of household utility). However, the household’s decision about how much to spend depends on the prices of the consumed goods, on the level of utility, and on other household characteristics (such as household composition).

A household’s income, on the other hand, comprises the sum of the wages of all working members of the household plus profits from business or other economic activities. Profits include, for instance, the net income from agricultural production or farm enterprises; they depend on prices, technical capacity, and household characteristics. Since profits are defined as sales net of purchases of inputs, some of the costs imposed by taxes on inputs or intermediate goods can be assessed by their impacts on profits. Income may also comprise transfers (public or private), savings, and other unmeasured returns to factors of production, such as land holdings.

This simple framework highlights the fact that household welfare depends on market equilibrium variables such as prices and wages (that affect household choices) and also on household endowments. For example, household consumption depends on the prices of consumer goods, and household income depends on the type of labor endowment (skilled, unskilled), the market wage rate, and the prices of goods produced by economic enterprises run by members of the household. It follows that changes in commodity prices affect welfare directly via consumption and production decisions, and that these impacts are heterogeneous insofar as they depend on household choices and endowments. In addition, there are short-run impacts when households do not adjust their consumption, employment, or
production choices; medium-run impacts refer to the welfare (or consumption or income) effects that take into account partial adjustments; and long-run impacts take into account income growth, investments, and other factors that are related to long-run choices. In other words, the time horizon associated with the impact of price changes on household welfare is commonly defined by the extent to which a household has been able to optimize its total consumption or income in light of the price changes.

**First-order Impacts**

Let us consider now the impacts of changes in the price of a commodity. To simplify, we follow Deaton (1989a) and assume that the principle of “separability” holds. Under this assumption, production decisions are independent of consumption decisions (utility maximization). This means that we can consider the income level of the household as exogenous (once optimal production decisions have been made) when utility maximization takes place as the household adjusts its consumption bundle. The separability assumption is not just a technicality: it allows us to simplify the welfare analysis, but it is valid only under certain economic conditions, such as perfect and complete markets for goods, credit, insurance, and so on.

Prices affect both expenditures and income, even if they do so separately. On the consumption side, consumers are worse off if prices go up but are better off if prices go down. In a first-order approximation, these impacts can be measured with budget shares. That is, the size of the effect of a price change on a household directly depends on how important the good is in the household’s total expenditures. On the income side, there is also a direct impact on profits if the household produces the affected good, which depends on the share of income attributed to this good. In rural economies, this source of income can account for a large fraction of total income. In more urbanized economies with more developed labor markets (as in many places in Latin America), the role of the direct production of (agricultural) goods tends to be less important.

Comparisons of budget shares and income shares establish a key result in the literature: a price increase hurts net-consumers (defined by the difference between expenditures and incomes associated with the commodity whose price changed), and benefits net-producers. The opposite is true for price declines: net-consumers are made better off and net-producers worse off. Further, the direction of welfare impacts carries over to the level of a national economy or any aggregate unit of analysis: a country that is a net exporter of an agricultural good will benefit, on average, from price increases, whether they come from global market fluctuations or from international policy reforms such as the international liberalization of agricultural trade policies or subsidies; but net-importers will probably be hurt by those
changes. In the context of the recent food crisis, for instance, this implies that countries that are net-importers of food were in principle hurt, while net-exporters were the winners.

Households, Net-Consumers, Net-Producers and Inequality

The net-consumer/net-producer result was introduced by Deaton (1989a), who used data from the Thailand Socioeconomic Survey of 1981–82 to explore the distributional consequences of an export tax on rice across Thai households. He found that an increase in the price of rice (resulting from the elimination of the export tax) would benefit the average household. The average poor, as well as the average rich, would benefit little. The benefits for the poor are small because they tend to both consume and produce lots of rice, thus selling little. The average benefit to the rich is also small because while sellers are often large, there are few of them among the rich. The gains would be much higher in the middle of the income distribution, indicating that the middle class would gain the most from higher rice prices.¹¹

The ideas introduced in Deaton’s work have been, and still are, extensively utilized in the literature. Early examples include the following authors: Deaton (1989b), who reviews applications for Cote d’Ivoire, Indonesia, and Morocco; Budd (1993), who investigates food prices and rural welfare in Cote d’Ivoire; Benjamin and Deaton (1993), who study cocoa and coffee in Cote d’Ivoire; Barret and Dorosh (1996), who look at rice prices in Madagascar; and Sahn and Sarris (1991), who examine structural adjustments in several Sub-Saharan African countries. Deaton (1997) provides an account of the early use of these techniques in distributional analysis of pricing policies.

Recently, there has been a rebirth of this type of work, motivated by the food-price “crisis” associated with skyrocketing prices in global markets. Ivanic and Martin (2008) cover ten case studies using household survey data from Bolivia (2005), Cambodia (2003), Vietnam (1998 and 2004), Malawi (2004), Nicaragua (2001), Pakistan (1999), Peru (2003), and Zambia (1998). These authors find an overall poverty-increasing impact of higher food prices because in the sample, most poor households are net-consumers of food (both in urban and rural areas).

Wodon et al. (2008) apply this method to a dozen West and Central African countries, including Burkina Faso (2003), Ghana (2005–2006), Democratic Republic of Congo (2005), Gabon (2005), Guinea (2002–2003), Liberia (2007), Mali (2006), Niger (2005), Nigeria (2003–2004), Senegal (2006), Sierra Leone (2003), and Togo (2006). The authors focus on rice, wheat, maize, other cereals, milk, sugar, and vegetable oils and find that a 50 percent increase in prices for selected food items would cause an average increase in the share of the population in poverty of between 2.5 and 4.4 percentage points. The average impact would be
between 3.7 and 5.2 percentage points in urban areas, and between 2.2 and 4.1 points in rural areas. These authors also find significant heterogeneity across countries, as expected.

While the net-consumer/net-producer result is intuitive, it rests on two critical assumptions, namely that the first-order approximation is valid and that labor markets are unaffected by the price change (Deaton 1997). The first-order approximation is analytically useful, but it is valid conceptually only for small price changes. We say “conceptually,” because in principle budget and income shares could be used to measure the distributional effects of both small and large price changes, but the approximation error will be larger when price changes are large. In these cases, it is crucial to incorporate supply and demand responses, and we discuss these below. The role of labor markets is discussed next.

### Labor Markets

Price changes also affect wages; the mechanisms are simple. Consider, for instance, a small, open economy that takes commodity prices as given in international markets. Changes in relative product prices cause some sectors to expand and some others to contract. As a result, labor demand for different types of labor in different sectors can change, thereby affecting equilibrium wages. In the literature, these responses are captured by “wage-price” elasticities, which measure the proportional response of wages to price changes. These elasticities may vary across members of a household when they are endowed with different skills (unskilled, semi-skilled or skilled labor), or if they work in different sectors (industry premia). Clearly, if countries differ in technologies, endowments, or labor regulations, the responses of equilibrium wages to prices can be heterogeneous across countries as well.

The response of wages can generate first-order effects on household welfare. To account for these responses, the standard net-consumer/net-producer proposition needs to be modified. To illustrate, consider the extreme case where a farm household consumes a product but does not produce it. Instead, the household earns income from selling labor to neighboring farms. By omitting wages, this household would be mistakenly classified as a net-consumer and could thus be presumably hurt by a price increase. But if wages respond positively to prices, the final welfare effect may not necessarily entail a loss.

This type of wage response was studied in Ravallion (1990), who explored the conditions under which net consumers of food products in Bangladesh lose or gain in the face of increased food prices when rural wages adjust. Ravallion estimates low elasticities of agricultural wages to food prices and concludes that responses are unlikely to be strong enough to offset the short-run adverse distributional effects of higher food prices. The long-run estimates appear to be more favorable to the poor,
but it would nevertheless take around four years for any gains to materialize. Boyce and Ravallion (1991) take another look at this issue using newer data for Bangladesh. These authors set up a dynamic econometric model of agricultural wages and rice prices and find that increases in rice prices relative to the prices of manufactured goods have adverse effects on the real wages (measured per quantity of rice) in both the short- and long-run.

In a study of informal export barriers and poverty in Moldova, Porto (2005) amends the first-order approximation used by Deaton (1989a) to allow wages to adjust. In this setup, exporters produce agro-manufactured goods (i.e., wines or apple juice) for international markets. These exporters purchase primary products from the farmers (grapes, apples) and manufacture exports using, among other inputs, labor. Farmers are endowed with land and labor and allocate time (net of leisure) to the “farm” labor market, the home (land) plot to produce primary inputs (grapes) for sale, or to work in nearby farms. In this framework, Porto derives additional welfare impacts given by the direct impact on wages. Also, a given farm can either sell or purchase labor in the “farm-labor” market, and consequently, the estimation of the first-order welfare effects needs to include the additional impacts that arise due to changes in wage earnings or paid wages. Porto (2005) finds that increases in the prices of agro-manufactured exports such as wines (a major export sector in Moldova) have sizeable poverty-reducing impacts. Wages respond positively to export prices and this causes first-order gains that dominate both the consumption losses due to higher consumer prices and the profit losses due to higher wages paid to hired labor.

Porto (2006) studies a case where households consume products and sell labor in the market, but do not produce farm products. This author’s model applies well to urban, middle-income economies where farm activities are not relevant (and are not captured by the available household surveys, either). Porto explores the distributional consequences of MERCOSUR (a regional trade agreement between Argentina, Brazil, Paraguay, and Uruguay) in Argentina and finds welfare gains for the average poor and middle-income households (and negligible effects for the wealthiest households). This is because on top of gains from price reductions due to tariff cuts, there are changes in wages that favor unskilled workers over skilled workers, and unskilled workers are concentrated at the bottom of the income distribution.

Three recent studies are also worth mentioning. Ferreira, Fruttero, Leite, and Lucchetti (2011) examine the consequences of food price inflation in Brazil, a large food producer with a predominantly wage-earning agricultural labor force. These authors find large and negative consumption effects, which are also markedly regressive, as expected. However, they also uncover a positive and progressive income effect, particularly in rural areas. Thus, overall the Brazilian middle-income household suffered larger proportional losses than the very poor or the richer households. Jacoby (2013) reaches a similar conclusion in his study of food prices in India.
Specifically, once the wage gains are accounted for, he finds that rural households across the income spectrum actually benefit from higher agricultural commodity prices. Finally, Nicita, Olarreaga, and Porto (2014) examine the presence of a pro-poor bias in the trade protection of the following six Sub-Saharan African (SSA) countries: Burkina Faso, Cameroon, Côte d’Ivoire, Ethiopia, Gambia, and Madagascar. These authors’ model includes consumption effects, agricultural income effects, and labor income effects; they find that trade policies tend to be biased in favor of poor households because protection increases the agricultural prices of goods that are sold by African households, and this effect dominates the impacts on both expenditures and wages.

Spillovers

Broadly speaking, spillovers are the impacts of a change in market $i$ on the activity in market $j$. There are two types of spillovers: 1) production linkages occur when the expansion of a sector affects upstream activities (backward linkages) or downstream activities (forward linkages); 2) expenditure linkages occur when the increase in income due to the expansion of a sector raises the demand for outputs and thus the derived demand for inputs in other sectors. Given our focus on commodities, we emphasize here spillovers in agriculture. The large body of literature on the topic uncovers two main regularities. First, the linkages are strong in rural areas because households earn a large share of their income from and spend a large share of their total expenditures on goods produced locally. Second, both production and expenditures spillovers are important (Mellor and Johnston 1984; Delgado 1996; Harriss 1987). Evidence on spillovers in agriculture in Latin America can be found in the regional study by De Ferranti et al. (2005). These authors report that while the share of the primary sector in GDP is 12 percent, there are massive spillovers to the rest of the economy. In fact, the multiplier factor from the primary sector to national GDP is 0.22, so that, for instance, a one percent growth of the rural natural resource sector is associated with an increase of 0.22 percent increase in national GDP (much higher than the 0.12 percent increase to be expected from the sector’s share of GDP). At 0.28, the multiplier for the income of the poorest households is even higher.

Note that spillovers can be interpreted as general equilibrium effects, but they are not necessarily the same. If there is some sort of labor immobility in general equilibrium, for instance, then a change in the price of good $i$ may have no effect on the wages paid in sector $j$. In such a scenario there are no spillovers, even in a general equilibrium setting. This distinction is useful to organize our discussion. Of the papers reviewed above, Ravallion (1990) estimates a partial equilibrium model and assumes no spillovers or general equilibrium effects from rice prices to non-agricultural wages.
An alternative interpretation is that labor markets are segmented. In contrast, Porto (2006) and Nicita (2009) take a general equilibrium approach and assume that a change in a given price $p_i$ affects economy-wide wages. This assumes that labor markets are integrated and workers are freely mobile across sectors; an increase in wages in one sector spills over to other sectors.

Porto (2008) describes a variant of the spillover mechanism. Markets can be segmented so that wages can differ across sectors. However, sectors are related via forward and backward linkages, so an expansion of one sector may have implications for other sectors. In such a setting, changes in corn prices in Mexico can affect the corn sector directly, as well as various rural services associated with corn production. For instance, higher corn prices may require more labor in corn planting, weeding, and harvesting (thus causing wages in the corn sector to increase), and may also require farm services (road graveling, farm-gate and fence repair, etc.). Porto finds that allowing for those spillovers can make a big difference in assessing the welfare impacts of higher corn prices. In particular, the average Mexican low-income household tends to be a net-consumer of corn and consequently is made worse off by an increase in corn prices. However, once spillovers from the corn sector to agricultural wages are allowed to take place, even low-income net-consumers stand to gain from higher corn prices.

As should be clear from this review, the existence or lack of spillovers, and the nature of those spillovers, will be specific to the case under study. It is a priori difficult to make generalizations.

Different types of spillovers arise when other product markets, rather than labor markets, are affected. These spillovers are likely to take place in non-traded goods. As we have shown above, changes in commodity prices affect factor prices, including wages. If the wages earned in non-traded sectors are affected, then the cost of producing these goods will change, and this in turn will affect the equilibrium prices of these goods. As a result, there are additional welfare impacts on the consumption side. Notice that these are first-order impacts. Porto (2006) provides an example of how to estimate these impacts for the case of the countries that are members of the MERCOSUR customs union; he uses time series of prices to recover the elasticity of the price of non-traded goods with respect to the prices of traded products; import tariff cuts by MERCOSUR appear to cause the prices of non-traded goods to decline and households to benefit. It is noteworthy that these spillovers are similar to those highlighted by, for instance, Burstein, Neves, and Rebelo (2003) for the computation of the real exchange rate and the adjustment of prices to stabilization programs.

These types of spillover can be particularly important for some commodity price shocks, especially on energy. Households do spend money directly on energy (electricity, heating), but sizeable effects can be created by spillovers to other good prices. For instance, oil price shocks can affect households via expenditures on transport;
the direct impacts can be calculated using (exogenous) price changes or shocks and budget shares, as carried out above. To measure the indirect impacts we need to estimate the response of the prices of non-traded goods after commodity price shocks. This can be done with the “price shifting” model proposed by Coady and Newhouse (2006) and Coady et al. (2006). In this methodology, commodity price shocks are fully transmitted to non-traded prices using input-output tables to infer the share of different (commodity) inputs in the production cost structure. The welfare effects at the household level can then be approximated by multiplying the induced price change of non-traded goods and the budget shares from the surveys.

This methodology has been applied extensively to study the distributional effects of fuel subsidies. Much of this work is summarized in Arze del Granado, Coady, and Gillingham (2012), who present cross-country welfare impacts of an increase in the retail price of fuel products by $0.25 per liter. These authors find substantial welfare effects of the total (direct and indirect) impact of such a price increase. On average, worldwide, household real income would decline by 5.4 percent; there is regional heterogeneity, as expected. The average welfare impact is the highest in Middle East and North Africa (7.4 percent), Africa (5.4 percent), Asia and Pacific (4.5 percent), and South and Central America (3.3 percent). A key finding of this line of work is that the spillovers—the indirect effects—are very important. On average, the indirect impact accounts for 61 percent of the total welfare effect.

Price Pass-through

So far, the review has focused on a generic analysis of price changes. Often these price changes, especially if they refer to commodities, are generated by external shocks. It is thus important to review the literature on the pass-through of international prices to the domestic economy, and furthermore to the household. There are various bodies of literature that are relevant here: international economics, international trade, and economic development. While this collection of theories and empirical analyses play different roles in our framework, we can learn from all three of them.

Standard models of international trade and international economics assume competitive markets (and homogeneous goods) and frictionless trade. In this scenario, markets are integrated and the law of one price holds: domestic prices are equal to the international prices converted to the local currency. Any difference between these prices is due to transport and distribution costs as well as to trade policy. In this family of models, a proportional change in the exchange rate, in the international price, or in the tariff is fully transmitted to domestic prices. There is, however, strong evidence against this prediction. The papers reviewed by Rogoff (1996) consistently reject the law of one price for a variety of products and countries.
The second type of impact arises when distribution costs change. Distribution activities are mostly non-tradable services, which require factors of production such as labor and land. If factor prices react to changes in international prices or exchange rates, then distribution costs will themselves change, and this increase can feed into a larger-than-proportional increase in the domestic price of imported goods. There is evidence that this mechanism can be relevant in practice. In Argentina, for instance, Burstein, Neves, and Rebelo (2003) show that the distribution margin in both wholesale and retail is around 60 percent. Using calibration methods to compute the impact of a stabilization program (i.e., they study a decline in the exchange rate), the authors find that the increase in the relative prices of non-tradable goods can be 60 percent higher in a model with distribution costs than in a model without such costs. With perfect pass-through, Burstein et al. (2003) assume that the prices of tradable goods decrease in the same proportion as the decrease in the exchange rate, and this affects the prices of non-tradable goods. Without distribution costs, the authors calibrate an increase in the relative price of non-tradable goods of 15.4 percent. With distribution costs, the increase in the price of non-tradables is 25.7 percent.

Another strand of literature is the work on exchange-rate pass-through (ERPT), which is defined as the percentage change in import prices resulting from a 1 percent change in the exchange rate. ERPT models are often associated with imperfect competition in product markets (but they are consistent with perfect competition as well). Goldberg and Knetter (1997) identify two conditions for full pass-through: constant mark-ups, and constant marginal costs. These conditions are unlikely to hold in a given market, and as a result, the literature has found strong evidence of imperfect pass-through. In their review, Goldberg and Knetter (1997) conclude that we should expect a pass-through rate of around 60 percent. Because this paper includes measures of costs in the derivation of the pass-through rate, the results imply that around 40 percent of the change in the exchange rate is actually offset by changes in markups. In a comprehensive analysis of pass-through rates, Campa and Goldberg (2005) find large differences across developed countries. The estimated pass-through rate is 42 percent in the United States, 98 percent in France, 80 percent in Germany, and 46 percent in the United Kingdom.

An interesting paper that attempts to incorporate the estimation of tariff-rate pass-through into the trade and poverty literature is that of Nicita (2009). This author argues that while imperfect price transmission has been established in the literature (as our review suggests), most of the work on the impacts of trade on poverty, and more generally on the distributional impacts of pricing policies, implicitly assume full pass-through. Nicita attempts to fill this gap by estimating tariff-rate pass-through for Mexico; he finds lower pass-through rates for both manufacturing (27 percent) and agriculture (33 percent), and supports these findings by arguing that pass-through is likely to be smaller in a developing country like Mexico due to a lack of domestic transport infrastructure.
Another important strand of the literature is the work on spatial market integration and the efficiency of arbitrage. In the presence of transport costs, prices across regions differ in the size of the transport costs but should move together due to arbitrage. In practice, three scenarios can be observed: efficiency, characterized by price differentials that can be perfectly accounted for by transport costs; segmented efficiency, characterized by price differentials that are smaller than transport costs (so that markets are competitive but segmented and high transfer costs make trade and arbitrage not profitable); and inefficiency, characterized by price differentials that are larger than transfer costs and thus imply unexploited arbitrage profits (due to market power, for instance). This is relevant for our review because only if markets are efficient will perfect pass-through take place.

There are various procedures to test for spatial market efficiency. The state-of-the-art method is the parity-bound model of Sexton, Kling, and Karman (1991), Baulch (1997), and Park et al. (2002), among others. The method relies on maximum likelihood methods to estimate a switching regression model for the three states, namely efficiency, segmented efficiency, and inefficiency. As expected, the evidence is mixed. Baulch (1997) finds that rice markets in the Philippines are spatially efficient and that differences in prices levels are due to transaction costs. Instead, Cirera and Arndt (2006) report inefficiencies in maize markets in Mozambique. Notably, these inefficiencies are somewhat reduced by market liberalization (but not by road rehabilitation), but tend to persist over time. This indicates imperfection in the pass-through rate of commodity shocks.

Household Adjustment in Consumption and Production

The discussion has followed the traditional first-order approach to assessing the distributional effects of commodity price changes. A major concern with this approach is that it does not allow households to respond to the changes in the economic environment. In fact, even when wage adjustments are incorporated, the standard approach takes the net-consumer/net-producer position of the household as “exogenous.” This is just a datum in the analysis that we can assess using survey data, as we did when we described the dependency of households in Latin America and Africa. The net-position of the household is, however, endogenous when households respond.

It is often argued that the first-order approximation can be accurate for small price changes. The “endogeneity” of a household choice in these cases is a second-order concern in the sense of being unimportant. This is true, and there are numerous examples where the framework works very well. The evaluation of the Doha Development Agenda is one instance, as explored by Hertel and Winters (2006) or Hoekman and Olarreaga (2007). For larger price changes, such as those that can...
be expected from a persistent food-price crisis (or a collapse of global commodity prices), the approximation error grows larger. This section thus explores how large this error can actually become and discusses how the approximation can be improved by measuring household responses in consumption and production.

A consumer will always lose from an increase in prices, but the losses can only be ameliorated by reducing purchases of the more expensive goods. A producer will instead always lose from a price decline. In this case, adjustments can mitigate the losses by shifting resources to more productive activities. Similarly, a farm that hires labor will lose from higher wages, even if labor is substituted with other production factors. Finally, a worker will lose from a reduction in wages, even if he decreases labor supply to enjoy more leisure.

For a farm-household, or more generally in a model where the household participates in multiple markets, adjustments can play a bigger role. In fact, Porto (2008) shows that it is possible for a household to change sides of the market (that is, to become a net-producer after a sufficiently high price increase) so that even an ex ante net-consumer can benefit from a price increase. In what follows, we illustrate various sources of adjustment and discuss attempts to quantify them.

Consumption Responses

We begin with consumption because research in this area has been prolific. To incorporate consumption responses, research needs to estimate a system of demand elasticities (own- and cross-price elasticities) to identify the pattern of substitution in consumption. The estimation of such a demand system is difficult, however.

There is a long tradition in economics in estimating systems of demand equations. The type of estimation method depends critically on the available data. In the early days when only national account data were available, estimation relied on time series of various consumption aggregates and price indices. The search for a utility-consistent empirical model of demand converged to the almost ideal demand system (AIDS) model of Deaton and Muellbauer (1980), which is now the leading framework for studying demand elasticities.

The mass-production of household surveys in more recent years—a process in which the World Bank has been instrumental—has broadened the options for demand analysis. However, while expenditure surveys provide detailed accounts of what people consume, they collect little information on prices. In some cases, community questionnaires collect price data at the village level, and this variation can be exploited to identify demand elasticities. For those cases in which the surveys collect information on quantities purchased as well as on expenditures, demand elasticities can be recovered from unit values—the ratio of expenditures over quantity. Deaton (1987, 1988, 1990) developed a model of demand for quantity and
quality that allows for the extraction of the price signal embedded in the unit values at the level of a market (or cluster of households residing near each other, and thus shopping for consumption goods in the same markets), and these price signals identify the demand elasticities. Because the estimating equations are not necessarily derived from a utility framework (that is, the equations do not form a full-blown AIDS model), the estimated parameters cannot be used to recover the parameters of the utility function. However, they do reveal how household consumption responds to price changes.

Deaton’s unit values procedure has been used extensively to examine issues of tax reforms (see, for instance, Deaton (1989b), Deaton and Grimard (1992), and Deaton (1997) for case studies of Thailand, Pakistan, India, and Cote d’Ivoire). A nice application of Deaton’s methods to welfare analysis is the research conducted by Friedman and Levinsohn (2002), who studied consumption responses during the Indonesian financial crisis of 1997. These authors compared the welfare effects resulting from a first-order approximation (that is, without substitution effects) with the ones resulting from a second-order approximation (thus including substitution effects). However, they did not account for wages or, more generally, income responses. Two major findings are worth emphasizing in the context of the discussion about the magnitude of consumption responses to market signals. On the one hand, the losses estimated as first-order approximations for the urban poor were high at roughly 50 percent of the pre-crisis level of per capita expenditures. In rural areas, the losses of the poor were equivalent to nearly 40 percent of their pre-crisis real consumption. On the other hand, allowing for substitution effects in consumption made a big difference. The estimated losses were cut by half for the urban poor and by more than 40 percent overall. These results imply that first-order approximations can be quite large when price shocks are large, such as those observed during the Indonesian crisis (when prices increased by 100 percent or more). In cases of economic crisis, when large price changes may be the norm rather than the exception, the results of Friedman and Levinsohn (2002) warn us about the perils of ignoring second-order terms in welfare analysis. In light of the stylized facts presented above, these adjustments need to be taken into account in the context of the recent surge of commodity prices.

**Farm Supply Responses**

We turn now to farm supply responses. As throughout all of our analysis, we are interested in responses at the household level, and especially for the poorest households. Since these households are typically involved in agriculture, much of our review is concerned with agricultural farm responses. Conceptually, the issues are those that we discussed within consumption responses. Faced with changed relative prices, the question is whether there is any adjustment in
production and what factors prevent or facilitate that adjustment. It is perhaps not really surprising to learn that supply responses are somewhat feeble in developing countries. The body of literature is too large for a comprehensive review, but some of most fascinating recent examples include Cadot, Dutoit, and de Melo (2009), who examined vanilla-market reforms in Madagascar, and McMillan, Welch, and Rodrik (2003), who studied cashews in Mozambique.

The evidence on supply responses is not always so bleak, however, as two strands of literature illustrate. First, a common finding in empirical work is that sizeable responses can be expected when price changes are accompanied by appropriate complementary factors such as access to inputs, information, and credit (McKay, Morrissey, and Vaillant 1997), productive assets like animals and tools (Deininger and Olinto 2000), productive capital, education, and land quality (López, Nash, and Stanton 1995; Heltberg and Tarp 2002). Second, there is a growing body of literature showing responses to marketing reforms under successful contract farming arrangements. Two examples illustrate the point: Brambilla and Porto (2011) estimate large adjustments in cotton production after changes in the outgrower contract environment in Zambia, and Maertens and Swinnen (2009) estimate similarly large responses in the vegetable sector in Senegal.

Another way to explore these issues is to look at the performance of subsistence farmers relative to commercial farmers. Subsistence farmers produce for home consumption, whereas commercial farmers earn their income by selling their produce in markets. The evidence indicates that commercial farming is associated with higher consumption and incomes than subsistence farming. Cadot, Dutoit, and Olarreaga (2006) estimate an average loss in income of those in subsistence farming of 43 percent relative to commercial farmers that produce the same crop. Even with much smaller margins, what is puzzling is why subsistence farmers do not transform themselves into commercial farmers by selling their harvest on the market? This question is essentially about supply responses to market signals.

Cadot, Dutoit, and Olarreaga (2010) argue that there are barriers to exit from subsistence agriculture, but the issue highlights explanations of why supply responses to market signals might be sluggish. There are two major barriers here. One is risk: cash crops are riskier than food crops so that part of the difference in return is actually a risk premium (Fafchamps 1992; Rosenzweig andBinswanger 1993). The other barrier is related to missing markets: the differential rate of return is due in part to the shadow price of a good with missing markets such as for labor or food crops (de Janvry, Fafchamps, and Sadoulet 1991). Notice that missing markets can arise due to variable transaction and transportation costs, fixed transaction costs (searching for partners, enforcing contracts with distant buyers, establishing quality), or sunk costs of transacting. One example is Key, de Janvry, and Sadoulet (2000), who illustrate how supply responses are affected by transaction costs. Another interesting example is Cadot, Dutoit, and Olarreaga (2006), who
show that sunk costs can be particularly large, ranging from 124 to 153 percent of
the market value of the average annual farm output in Madagascar.

We end this section with a caveat. Positive and large supply responses to price
changes are always an indication that economic incentives work. However, in the
context of our analytical framework above, these supply responses comprise second-
order effects. In consequence, they are likely to generate only small welfare impacts
because of the opportunity cost of the resources used in the production of different
goods. While this argument does not invalidate the search for supply responses in
the data, it strongly argues for a careful interpretation of the findings.

Quantifying The Mechanisms: The Case of Mexico

In this section we use data from Mexico to quantify the different channels unco-
vered by our review of the literature. The data come from the 2006 Household
Income and Expenditure National Survey, ENIGH (Encuesta Nacional de Ingresos y
Gastos de los Hogares). The main objective is to illustrate how the mechanisms
work. To do this, we work with a (hypothetical) increase of 20 percent in corn
prices in rural Mexico. Our results are reported in table 4.

We begin with the first-order consumption and production effects and the net-
consumer-net-producer proposition. As discussed in our framework above, the con-
sumption effects are calculated as the product of the budget share spent on corn and
the 20 percent price increase. The production effects are the product of the income

Table 4. Welfare Effects of Price Changes, Examples from Mexico

<table>
<thead>
<tr>
<th></th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>All</th>
<th>Net Producers</th>
<th>Net Consumers</th>
</tr>
</thead>
<tbody>
<tr>
<td>First-order Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumption effects</td>
<td>−3.56</td>
<td>−2.53</td>
<td>−1.99</td>
<td>−1.40</td>
<td>−0.73</td>
<td>−1.99</td>
<td>−0.99</td>
<td>−2.70</td>
</tr>
<tr>
<td>Production effects</td>
<td>2.25</td>
<td>1.86</td>
<td>1.65</td>
<td>1.49</td>
<td>1.12</td>
<td>1.60</td>
<td>2.78</td>
<td>0.77</td>
</tr>
<tr>
<td>Net effects</td>
<td>−1.31</td>
<td>−0.67</td>
<td>−0.34</td>
<td>0.09</td>
<td>0.38</td>
<td>−0.39</td>
<td>1.78</td>
<td>−1.93</td>
</tr>
<tr>
<td>Segmented Labor Markets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income effects</td>
<td>5.86</td>
<td>4.84</td>
<td>4.28</td>
<td>3.87</td>
<td>2.90</td>
<td>4.16</td>
<td>7.22</td>
<td>1.99</td>
</tr>
<tr>
<td>Net effects</td>
<td>2.30</td>
<td>2.31</td>
<td>2.29</td>
<td>2.47</td>
<td>2.17</td>
<td>2.17</td>
<td>6.23</td>
<td>−0.71</td>
</tr>
<tr>
<td>Labor Market Spillovers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income effects</td>
<td>8.67</td>
<td>7.21</td>
<td>6.45</td>
<td>5.86</td>
<td>4.78</td>
<td>6.31</td>
<td>10.77</td>
<td>3.14</td>
</tr>
<tr>
<td>Net effects</td>
<td>5.10</td>
<td>4.68</td>
<td>4.47</td>
<td>4.46</td>
<td>4.04</td>
<td>4.32</td>
<td>9.78</td>
<td>0.44</td>
</tr>
<tr>
<td>Second-order Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumption effects</td>
<td>−3.34</td>
<td>−2.37</td>
<td>−1.87</td>
<td>−1.31</td>
<td>−0.69</td>
<td>−1.8</td>
<td>−0.93</td>
<td>−2.53</td>
</tr>
<tr>
<td>Production effects</td>
<td>2.48</td>
<td>2.05</td>
<td>1.81</td>
<td>1.64</td>
<td>1.23</td>
<td>1.76</td>
<td>3.05</td>
<td>0.84</td>
</tr>
<tr>
<td>Net effects</td>
<td>−0.87</td>
<td>−0.33</td>
<td>−0.06</td>
<td>0.32</td>
<td>0.54</td>
<td>−0.11</td>
<td>2.12</td>
<td>−1.69</td>
</tr>
</tbody>
</table>

Note: Authors’ own calculations based on Porto (2008).
share of corn production and the 20 percent price increase. The net effect is the difference between the production share and the consumption share. For the average Mexican household, an increase in corn prices brings only small welfare losses (equivalent to 0.39 percent of average national income). As expected, there is heterogeneity across quintiles. The largest losses are at the first quintile (−1.31 percent), while quintiles 4 and 5 instead slightly benefit from higher prices. In fact, for these households, the average production effect is larger than the average consumption effect so that poorer households tend to be net consumers of corn, while richer households tend to be net-producers of corn. This illustrates the net-producer/net-consumer proposition. We can also explore this directly by dividing households into net-producers (those that produce more corn than they consume, independent of the quintile) and net-consumers. The average net producer would enjoy gains equivalent to 1.78 percent of their initial (i.e., pre-price increase) average expenditure. Net-consumers would instead suffer a loss of 1.93 percent.

To incorporate labor market responses, recall that the impacts depend on how wages respond to price changes, and this in turn depends on whether we think that labor markets are integrated or segmented, and upon whether there are spillovers and backward or forward linkages (Ravallion 1990; Porto 2005; Porto 2006). Continuing with the case of corn in Mexico, we provide two examples in table 4. In the second panel, we assume that labor markets are “segmented”, and thus we only allow wages in the agricultural sector to respond to corn prices. In this case, we use an agricultural wage elasticity with respect to corn prices of 0.40 (as estimated in Porto 2008). Consequently, an increase in corn prices of 20 percent translates into (roughly) 8 percent higher wages in agriculture. This, in turn, represents a wage-income effect on household welfare that depends on the share of total household income accounted for by agricultural wages. It should be pointed out that we only have information on wage income, and not on labor expenses. Thus, the results from this experiment may be biased upwards. The evidence in Porto (2005) for the Moldovan case suggests that these biases are, however, not very large.

For the average Mexican household, allowing for wage effects induces positive overall welfare effects: higher corn prices would bring welfare increases of 2.17 percent. The average income effects (which includes both the direct production effects of 1.60 and the wage effect) would be 4.16 percent. Since the poor are also agricultural workers, the welfare losses observed above for quintiles 1 to 3 are turned into average welfare gains due to the additional impacts on wages. In fact, the net effect in this (hypothetical) scenario would be roughly similar across the entire income distribution. The income gains of net-producers jump to 7.22 percent and their net gain becomes 6.23 percent—nearly 3.5 times higher than before. For net-consumers the income gains are more modest at around 1.99 percent, and these gains are not enough to offset the consumption losses. In the end, even with wage responses, net-consumers would be made worse off by the rise in corn prices.
In another example, in the third panel of table 4, we allow for spillovers from corn prices to the wages of self-employed individuals (in rural areas). The idea is that increases in agricultural prices may increase derived demand for labor in services, odd-jobs and, more generally, in local rural labor markets (not exclusively in agriculture). For the calculations, we use the same wage-price elasticity of 0.40 as before (Porto, 2008). The welfare effects of this expanded wage-income channel thus capture the increase in wages of 8 percent (20 percent price increase times the 0.4 wage-price elasticity) filtered through the share of wage income in total household income. We estimate the following welfare impacts: the average national gain would be equivalent to 4.32 percent on initial income; poorer households benefit relatively more than richer households; the income gain of net-producers would be equivalent to 10.22 percent of their initial income, and the net gain would be 9.78 percent; net-consumers would also have an income gain of 3.44 percent and a net gain of 0.44 percent.

As argued above, the net position of the household can be endogenous and it can thus change in response to a price shock. Consider, for example, an increase in the price of corn and let households be able to adjust. If consumption responses are sufficiently strong (because households can substitute easily to other food products) and/or if supply responses are large enough, then a net consumer can in principle become a net producer, and thus benefit from the increase in corn prices. The standard argument against this caveat is that second-order adjustments of the type described here are often small and thus unlikely to change the implications of the first-order calculations. An example is shown in the fourth panel of table 4, where we estimate consumption substitutions and supply responses. To do this, we use an own-price consumption elasticity of corn of −0.61 (as estimated by Porto 2008) and we assume a unitary corn supply elasticity. To simplify the exposition, we ignore here the wage impacts. As the results show, allowing for consumption and production second-order effects does not, in fact, affect the results. The gains for net-producers are slightly larger and the losses for net-consumers are slightly smaller, but the welfare impacts are not affected much.

Conclusions

This paper provided an overview of the impacts of price changes on household welfare. We began by exploring commodity dependence in the data and by looking at expenditure and income dependence in the household surveys from Latin American and Africa. Our analysis suggests a strong dependence on commodity prices, a dependence that is biased towards the poor, especially on the consumption side.

The paper then addressed the conceptual issues related to commodity dependence. It did so by laying out a theoretical framework, which described how
changes in commodity prices affect consumption, production, wages, and household adjustments. Each of these aspects of the analysis of commodity dependence is illustrated with evidence from the existing literature.

Our review tells the story of heterogeneous responses to commodity price changes. Most shocks to commodity prices are international in nature and the pass-through to domestic prices is imperfect (that is, changes in international prices are not translated one-to-one to domestic prices). The pass-through is also heterogeneous across countries and depends on infrastructure, institutions, and market structure (the nature of imperfect competition in domestic markets).

Faced with a price shock, households are affected on both the consumption and income sides. On the consumption side, higher prices cause welfare losses. In Latin America and particularly in Africa, higher food prices can cause large welfare losses, especially for the poor. On the income side, rising commodity prices can bring gains or losses, depending on which goods are produced and also on how wages respond to those price changes. The evidence shows that households tend to earn a significant share of their total income from commodities, so the consumption losses can be ameliorated, sometimes to a large extent. This is most likely to occur when wages and spillovers from commodities to other sectors are taken into account, particularly if wages respond positively to increases in labor demand in primary sectors, and if the spillovers to non-primary sectors are positive (via either production or expenditure linkages). In the end, the overall welfare impacts depend on the net-consumer-net-producer position of the household, after taking into account wage income. Net-consumers will be hurt by price increases, and net-producers will benefit. As expected, these impacts are heterogeneous across countries. Some countries will gain and some others will lose. Also, in a country that gains from higher prices, there will be households that gain and households that lose out. Conversely, there will be winners and losers within countries that lose as an aggregate.

Household adjustments can work as a cushion protecting against price changes. Given a price increase in commodities, consumption responses can ameliorate consumption losses and production responses can boost income gains. Typical adjustments under the separability assumption appear not to be enough to overturn the initial impacts derived from a static view of commodity dependence. Non-separability, which is likely to be a real feature in many developing countries, can work in both directions. Sometimes, market failures can shut down markets completely so that price changes can become irrelevant for many households. While this can work as a protective device, it can also hinder the realization of any efficiency gains and of expanded market opportunities. Sufficiently large price changes in the presence of market failures can generate discontinuities in household responses that can potentially multiply the typically small responses found in the empirical literature—in the real world, the price is not always right.
We end with some caveats. First, due to space limitations, our review has focused on welfare impacts and we have neglected issues of insurance, coping mechanisms, and social protection policies. Second, the review centered the discussion on one-shot price shocks. In particular, we have not covered the literature on the effects of commodity price volatility. In addition, even in the case of one-shot price shocks, we should note that there might be hysteresis in the effects. In other words, a price shock may have long-term implications, even if the price change reverts at a later date (Baird, Friedman, and Schady 2011). These themes require careful consideration in future assessments of the developmental implications of price signals.

Notes

Laura Chioda, Louise Cord, Irene Brambilla, and Emily Sinnott provided insightful comments during the early stages of this research project. We are grateful to Emmanuel Jimenez and three anonymous referees for detailed comments on a previous version of this article. We are especially grateful to Cecilia Peluffo and German Bet for their excellent research assistance. All remaining errors are our responsibility.

1. World Bank. Email: dlederman@worldbank.org.
2. Universidad Nacional de La Plata, Calle 6 e 47 y 48, Oficina 521, La Plata 1900, Argentina. Email: guido.porto@depeco.econo.unlp.edu.ar.
5. See, for example, “Life after the Commodity Boom,” The Economist, March 29, 2014.
7. Let $s_{it}$ be the commodity income share of household $i$ at time $t$. The cross-section model for $t$ is $s_{it} = x_{it}^0 \hat{\beta}_t + e_{it}$. For $t+1$, the model for household $j$ is $s_{j,t+1} = x_{j,t+1}^0 \hat{\beta}_{t+1} + e_{j,t+1}$, where we use $j$ instead of $i$ to highlight the fact that these are repeated cross-sections, not panels. We denote $\hat{\beta}_{t+1}$ the OLS estimates of $\beta_{t+1}$. From the empirical distribution of the residuals, let $\hat{e}_{j,t+1}$ be the draw attached to household $i$. With these estimates, the prediction for the income commodity shares for household $j$ at time $t$ is $s_{j,t+1} = x_{j,t+1}^0 \hat{\beta}_{t+1} + \hat{e}_{j,t+1}$. Note that, since we are taking averages, the residuals vanish (or are very small). Consequently, the lower and upper bounds in Dang, Lanjouw, Luoto, and McKenzie (2011) are very close in our experiment. Below, we report the lower boundary.
8. An alternative method is to start with the indirect utility function (see Deaton 1997).
10. To further simplify the exposition, we also assume separability between consumption and leisure in utility.
11. Notice, however, that this analysis does not take into account the fiscal implications of eliminating the tax.
12. Notice that we are referring here to prices of tradable goods. For non-traded goods, the issues are more complex since they are endogenous.
13. Note that this is a type of spillover as those discussed above.
15. In short, the difference is that compensated demand elasticities are needed for welfare analysis, and uncompensated elasticities are needed to analyze tax reforms.
16. We do not cover aggregate supply responses in our review.
17. We thank a referee for pointing this out.
18. The formulas for the calculations of the second-order effects are in Porto (2008).

References


194


