Procurement of Consulting Services for Implementation of a Prospective Module about the Informal Economy in the IFs Model, in order to Adapt it to the Case of Peru

(Contract 13-2014-CEPLAN)

Final Report to CEPLAN from the
Frederick S. Pardee Center for International Futures
Josef Korbel School of International Studies
University of Denver
www.pardee.du.edu

JULY 2015
INFORMAL ECONOMY REPORT V24.DOCX

David Bohl, Barry Hughes,
Mohammod Irfan, Eli Margolese-Malin, and José Solórzano

Support also from John McPhee, Peter Oesterling, and Caleb Petry

INTERNATIONAL FUTURES AT THE PARDEE CENTER
EXPLORE, UNDERSTAND, SHAPE

Table of Contents

Executive Summary **4**

1. Introduction: Motivation and Intent **9**

2. Conceptualization and Data **10**
What is informality? *10*
Characterizing informal economic activity *13*
The non-observed economy (UNECE 2008; Gyomai 2012) 13
The Shadow Economy (Schneider and Enste 2000) 15
Characterizing informal employment (ILO and WIEGO) *16*
Our own conceptualization *19*

3. Moving from Conceptualization to Measurement **22**
Estimations of informality *22*
Direct "survey" methods 23
Indirect "indicator" methods 23
Model-based approaches 24
The non-observed economy (NOE) method and results 27
Measuring the size and dimensions of the informal labor force 32
Comparing measures of the informal economy *37*

4. Drivers **41**
What does the literature say about what drives the size of the informal economy? *41*
Survivalist economies and the transformative power of deep developmental drivers. 43
Drivers of entrepreneurial informality 46
Push factors from formality to informality: Taxation, regulation, and corruption 46
Pull factors from informality to formality: Government transfers to households and spending on R&D 48
Drivers in IFs *50*
The variables analyzed for inclusion in the IFs model of informality 51
Developmental driver 51
Push factors 52
Pull factors 53
Structural Variables 54
Statistical summary 55
Building the models 58
Informal labor as a percent of total labor 58
Informal labor residing inside the informal sector 62
Informal GDP share: NOE and Shadow Economy 63
Historical analysis of informality models *65*

5. Impacts of Economic Informality **69**
Informality's linkage to productivity and economic growth *70*
Informality's linkage to taxes, social provision, and inequality *72*

6. Structure of the Informal Economy Representation in IFs **76**
Forecasting the informal economy *76*
Elaboration of core elements in forecasting the informal economy *77*
Forecasting informal labor share 78
Forecasting informal labor share by sector 81
Forecasting informal GDP share 82
Forward linkages: general approach *85*

Forward linkages: productivity and GDP 86
Forward linkages: government revenue 89

7. Analysis 91
Base Case analysis 91
Counterfactual: decreasing the informal GDP share and examining impact 94
Drivers of informality and the impact they might have 97
The individual interventions 97
Combined interventions and conclusion 101

8. Conclusions 103

Bibliography 104
Works Cited 104
Additional Sources Consulted 110

Appendix A1: Summary of Informal Economy Conceptualizations 113

Appendix A2: Summary of Methods for Estimating Informality 116

Appendix A3: Definitions of Informal Sector Enterprises and Labor used by selected countries 119

Appendix A4: Statistical Justification for Model Specification 121
Informal labor as a percent of total labor 121
Informal labor residing inside the informal sector 124
Informal GDP share (NOE formulation) and shadow economy 125

Appendix A5: Informal Economy Historical Datasets Added to the IFs Model 129

Executive Summary

Although the existence of informal economies in countries around the world has both benefits and costs for societies, the analytical consensus is that costs generally outweigh the benefits. Yet in such analysis, even conceptualization and measurement of the informal economy is challenging. And we are unaware of any significant quantitative studies that forecast the size of the informal share of labor forces and GDP in coming years, much less of attempts to dynamically explore the implications for countries of alternative development patterns with respect to that size. This study thus very substantially breaks new ground.

We begin by reviewing the use of terms including hidden, grey, shadow, underground and parallel, as well as informal economy, and explaining some of the variation and ambiguity of meaning in the concepts. We define our primary focus as clearly being on the informal economy, generally consistent with the Schneider (2013: 25) definition of "...market-based production of legal goods and services that are deliberately concealed from public authorities...." We do, however, focus also on another variable, the shadow economy, tied to data and estimations that imply a broader conceptualization than that of the informal economy. Another conceptual decision is whether to focus on informality in terms of the share of labor or the share of GDP that is informal. We have gathered data on and forecasted both informal labor and GDP shares.

Moving to data, we review data generated in three different ways, namely direct methods such as surveys or audits, indirect methods such as the use of proxies, and model-based methods. When possible we have looked to data from those using the direct methods. While adding multiple series to the database of the International Futures (IFs) forecasting system, for the purposes of initializing IFs we have used for labor a series from a collaboration of the International Labor Organization (ILO) and Women in Informal Employment: Globalizing and Organizing (WIEGO), blending that with another from the World Bank's Gender Statistics database. For informal GDP share, we turn to the non-observed economy (NOE) adjustments to national accounts as reported in the United Nations Economic Commission for Europe's 2008 survey of country practices and used by the Organization for Economic Cooperation and Development (OECD). In order to obtain more extensive country coverage, we adjusted model-based estimates from Schneider et al. (2010) and Elgin and Oztunali (2012) to NOE levels and blended it into the NOE series. An important insight from the two resultant series is that the share of informal labor in most countries significantly exceeds the share of informal GDP, reinforcing the likelihood that the informal economy is substantially less productive than the formal one. Although we believe that the NOE data, based on surveys, are stronger than the "shadow economy" data from Schneider et al. (2010) and Elgin and Oztunali (2012), based as they are on model calculations, we also use the shadow economy estimates to initialize an alternative variable in IFs for forecasting the size of the shadow economy.

Our analysis of drivers of the size of the informal economy (both labor and GDP shares) casts a very wide net; see the bibliography of this report. Most of that literature focuses on factors that keep enterprises out of the formal economy, often as a result of government inaction, such as failure to protect property rights, to eliminate inefficiency in registration systems, and to control corruption, or of government action, such as high regulatory and taxation burdens. There is little literature on forces that pull enterprises into the formal economy. In addition to our literature review, however, we undertook an extensive empirical analysis ourselves, statistically examining the explanatory power of a large number of possible driver variables from the literature and our own identification.

One of the striking features of our empirical cross-country analysis is the very strong relationship across countries between the size of the informal economy in terms of either labor and GDP shares, on the one hand, and deep and generally quite slowly changing developmental variables such as GDP per capita and the formal educational level of adults, on the other hand. These deep drivers are not highlighted in much of the literature, in part because analysis often has had relatively short-term temporal focus. Yet those variables are critical to us because of our long-term forecasting horizon. Our empirical analysis also identified variables with strong explanatory power among more proximate and potentially somewhat faster changing (and often policy-relevant) variables, such as the business climate, tax rates, corruption level, the extent of social support systems, and investments in economic restructuring such as R&D expenditures.

Our work built a model of the informal economy that is fully integrated with the larger International Futures (IFs) system. We focused on the drivers of labor informality and then, to maintain logical consistency between labor and GDP shares, we drove GDP share in part from the labor share, adding attention also to corruption and to research and development.[1] From our extensive literature and empirical analysis we identified three somewhat discrete clusters of drivers of labor share. At the deep socio-economic level we picked adult education years over GDP per capita; the two drivers are highly correlated, so forecasting should use one or the other, and education is both stronger statistically and an important developmental policy focus in countries attempting to navigate the middle-income passage to high-income status; in contrast, GDP per capita is more a result than a lever. At the more proximate level we conceptually divided variables roughly into push factors that tend to encourage informality (such as business climate and tax rates) and pull factors that tend to encourage formality if they are strong enough (such as social support spending and R&D). Although conceptually useful, that push

[1] In our analysis of the shadow economy we found that the same variables produced the best model (with, of course, different coefficients).

and pull division is somewhat blurred and has negligible implications for the model formulation.

The final model for informal labor share included years of adult formal education, an index of government business regulation (or inversely of friendliness), the tax rate on firms, and the extent of government transfers to households (social support and pension payments). Our final model for informal GDP share included the informal labor share, which has a strong relationship, the level of corruption using the Transparency International measure, and the rate of combined private and public spending on research and development (R&D) as a percentage of GDP.

Turning to the forward linkages or consequences of informality, we found that the literature on these is very sparse, although there is a general consensus that informality slows GDP growth. There are, however, some fundamental structural realities that suggest approaches to modeling two of those linkages, namely to productivity (and thus GDP and the broader model) and to governmental revenues (and thus government finance and expenditures). With respect to productivity, the data on which we rely suggest that the size of informal labor share in countries is almost always greater than the size of informal GDP share, up to about 3 times as great (for example, 60 percent informal labor versus 20 percent informal GDP). That implies an under-utilization of informal labor, and it indicates the rough magnitude of the potentially greater economic output were labor formalized (estimates of the "shadow economy" size are, however, much closer to those of the informal labor share, suggesting much less loss of productivity relative to that in the formal economy). Second, the taxation of the informal economy is almost near zero (with the exception of indirect taxes on purchases that are very difficult to avoid); thus formalization of economic output should create almost directly proportional government revenue-raising potential. We used both of these structural insights to build our formulations for forward linkages. The report text elaborates the model we created of both informal share drivers and forward impacts.

Finally, we undertook initial analysis of the model we built with a focus on Peru. That analysis has three elements. First we looked at the Base Case forecast, which shows the large size of the contemporary informal labor and GDP shares in the country (in 2010 roughly 70 percent for labor and 19 percent for informal GDP—and 53 percent for the shadow economy) and suggests the likelihood of substantial continued decline in those percentages, albeit over a long time period. Given a growing total workforce, the actual number of informal workers in Peru will fall only after a lag, but that decline should also begin in the 2020s.

The model includes parameters to adjust the Base Case pattern of change, allowing alternative scenarios, including potentially an alternative Base Case, for instance a forecast without informal share decline. We used the parameters to explore the behavior of the model in three ways. First, we asked counterfactually what might be the forward impacts if the informal GDP share were not to exist (to drop to 1 percent) between 2016 and 2025. That purely exploratory scenario generated an increase in Peruvian GDP of 15 percent by 2050 and a nearly 11 percentage point increase in government revenues as a portion of GDP. This purely hypothetical first scenario simply tests the forward linkages of informality reduction. We intended it to provide a ceiling in our analysis of potential benefits of informality reduction via policy leverage. It proved, however, as explained below, not to constitute such a ceiling.

Second, we began the process of exploring potentially viable alternative scenarios for reducing the informal labor (and thus GDP) share. In a sensitivity analysis, we individually manipulated our six drivers of labor and GDP share, allowing the model endogenously to pass forward changes in labor share to GDP share and then to productivity/GDP and government revenues. For instance, we lowered our five-variable government business regulation index by 50 percent over 20 years. This is a very significant, but not impossible change. More generally, we sought "aggressive but reasonable" levels for each of our sensitivity interventions. We found that such a business friendly posture had the greatest impact on informal labor and GDP shares of the variables in our model formulation.

We also found, however, that corruption reduction, followed by advance in adult education and higher R&D spending, had the greatest impacts on GDP growth through 2050. Why should they have greater impacts on economic growth than a more business-friendly posture, given that the latter reduces informality more rapidly? It is because those variables have additional impact in the IFs system on productivity and economic growth beyond those associated with the reduction in informality. That should not, of course, have been surprising, because changes in policy that reduce informality could easily also benefit the formal economy and, in fact, do so in the larger IFs forecasting system.

We should be careful in terms of assessing relative impact of interventions, because we need to recognize that the IFs system may not have fully represent all such secondary and tertiary impacts, including that of business regulation. Nonetheless, the result explains an analysis finding that emerged from our third analysis, combining the entire set of interventions. Specifically, the entire set boosts Peruvian GDP in 2050 by 26 percent relative to the Base Case, approaching twice the increase generated by the purely counterfactual scenario elimination of GDP informality after 2016.[2]

In consideration of this result, it is obvious that the set of six variables built into the formulations for forecasting informal labor and GDP share are, in fact, a package of policy intervention points that would not be altogether different, if somewhat less comprehensive, than a strategy for moving an economy through the middle-income

[2] Undertaking the same analysis with the shadow economy variable instead of the informal economy variable, the combined interventions generated a nearly identical 27 percent increase in GDP.

passage to high-income status—that is, for avoiding being caught in the so-called "middle-income trap."

Overall, this report outlines the building and initial use of the only analytical system that we believe exists for forecasting the future of the size of informal economies and exploring alternative scenarios. It suggests the very substantial benefits, which grow progressively larger over time, of national efforts to reduce informality.

1. Introduction: Motivation and Intent

Life in the informal economy has costs for individuals, enterprises, and society as a whole. For the individual, these costs include low wages and lack of social protections such as health care, job security, and pensions. For enterprises, which in the informal economy tend to be undercapitalized micro-firms or self-employed workers, the primary cost of informality is low productivity. There are also societal costs to informality. Because the individuals and enterprises of the informal economy generally escape taxation, the societal costs often include reductions in government revenues, meaning less expenditure capacity for everything from security to education and health to infrastructure provision. In addition, efforts to escape the scrutiny of and contributions to the public sector can lead to illegal behavior that feeds a pattern of corruption. Further, informal enterprises may avoid environmental and worker safety regulations.

At the same time, there are benefits to the informal economy. Informal activities can provide basic income for those unable to find employment in the formal economy and sometimes help them develop skills of use in formal positions. This can be especially important for countries facing a bulge of young and potentially under-educated people reaching working age. Thus, informal employment opportunities may even help maintain social stability. Firms may find that the costs of formality outweigh the benefits and can be economically more successful and more durable when informal. Parts of the informal sector can also be dynamically entrepreneurial, facilitating enterprises' and even the broader society's movement into new economic arenas.

Nonetheless, there is appropriate concern within societies with large informal sectors that the costs outweigh the benefits and that the society's movement out of low-income levels and through the middle-income passage to high-income status requires transforming informal activities to formal ones. Both the reality that high-income countries tend to have small informal economy shares and the analysis of development experts reinforces such concern.

This report and its associated modeling and forecasting, sponsored by Peru's Centro Nacional de Planeamiento Estratégico (CEPLAN), seeks to represent and analyze both the major forces or drivers behind informal economies and the implications or forward linkages of them to economic performance and government finances. We proceed by formulating our own conceptual framework for defining the informal economy based on the existing literature, exploring the availability of data by which to measure it, performing our own analysis of the drivers of its size, reviewing the literature on its implications for economic growth and government revenues, and documenting the development of our own computer model of the informal economy. We conclude by analyzing some results from use of that model.

2. Conceptualization and Data

What is informality?

There are many different terms for defining and classifying unregistered economic activities: hidden, grey, shadow, underground, and parallel, as well as informal. While these terms are often used interchangeably, they can carry with them varying definitions and connotations. For many years after its "discovery" in the 1970s (Hart 1973), the informal economy was considered more of a residual category, one comprised of the economic activities of urban, self-employed individuals that did not pay taxes, rather than its own economic sector (Gërxhani 2004). Subsequent attempts at a universal definition of the informal economy all fell short of capturing not just the multiplicity of activities, but the various political, economic, or social dimensions it includes. As Feige (1990: 6) suggests, "there is no single underground economy, there are many." Therefore, according to Gërxhani (2004: 270), "researchers gave up trying to formulate a unique definition, but instead, they have attempted to define the informal sector in accordance with the problem at hand." Contemporary definitions are thus inspired by a variety of motivations. As we will see, some analysts focus on the formalization of theory and rely on various quantitative tools and models to determine the drivers of informality, whereas others with more of a policy orientation frame their definitions within a typology that enables countries to better count the jobs and value added produced in the informal sector.

In surveying the literature, we find that there are four primary issues around which conceptualizations vary, even when the same label is used:

1. *Focus:* whether the focus of the study is on informal labor or informal enterprises (in the form of informal GDP), or both.
2. *The range of activities included*: is the conceptualization limited to otherwise legal activities that go unreported due to noncompliance with taxation and regulatory requirements, or are illegal activities also included?[3]
3. *The motivation for the informal activities:* a major division here is whether primary attention is on the inability of workers and/or firms to break into the formal economy, often for lack of adequate skills and capabilities (human capital), or whether dominant attention is on cost benefit decisions by actors around compliance with tax and regulatory requirements.
4. *Country differences*: whether there is a deliberate effort to capture distinctions in the nature of informality between developing and high-income countries.

[3] Illegal or illicit activities are most commonly defined as those comprising the production of goods and services whose sale, distribution, or possession is forbidden by law (e.g. illegal drugs), and those activities that are normally legal but are carried out by unauthorized producers (e.g. the unlicensed practice of medicine) (Eurostat 2013; SNA 2008).

Contemporary definitions of the informal economy have benefited greatly from the "the accumulation of rich data sets over the last decades," which "has cast progressively more light on the realm of the informal, permitting us to document the great heterogeneity of actors" and their reasons for operating in the shadows (Perry et al. 2007: 21). The proliferation of data has enabled researchers to develop detailed taxonomies of informality based on observable traits that allow for more comprehensive study of the sector's dynamics and diversity. Table 1 provides an overview of recent taxonomies and frameworks. As the table shows, most formulations of the informal economy include similar types of activities, especially underground productive activities (legal activities deliberately concealed from taxation and regulation) and informal production (legal activities by unincorporated and unregistered enterprises below a certain level of employment).[4] These activities are a subset of a larger informal ecosystem captured by some formulations that include illegal activities, household production for own-final use, and activities unreported due to statistical errors or shortcomings (Gyomai and van de Ven 2014; Eurostat 2013); other analysts do not make such extensions (Schneider 2012; Schneider and Este 2000).

Table 2.1. Summary of informal economy conceptualizations*

Source	Label/Focus	Definition	Components
Gyomai and van de Ven 2014; Eurostat 2013; UNECE 2008	Non-observed economy (NOE) Focus: adjusting formal GDP to capture informal GDP	Refers to all productive activities that may not be captured in the basic data sources used for compiling national accounts whether deliberately or through data deficiencies	**Underground production** (legal activities deliberately concealed); **illegal production**; **informal sector production** (activities by unincorporated/unregistered enterprises below a certain level of employment); **household production for own final use** (goods/services produced by the households that consume them); **and non-observed informal activities** (activities that should be accounted for but are missed due to deficiencies in the statistical programs)

[4] Several studies, including a World Bank report on informality in Peru arrange the various definitions of informality into two categories: (1) the productive definition, which defines the informal sector as economic units (typically the self-employed, family units, and micro-entrepreneurs and their employees) with low capital and low productivity; and (2) the legalistic definition, which considers all workers without rights to a pension (or certain other protections) to be informal workers, regardless of where they work (World Bank 2008: 19—20).

Elgin and Schneider 2013; Schneider 2012; Schneider and Este 2000	Shadow economy Focus: calculating size of informal GDP	All market-based legal production of goods and services that would contribute to officially calculated gross national product but are deliberately concealed from public authorities	Includes: **underground production; informal production** Excludes: **illegal activities**, **household own final use**
ILO 2013a; Eurostat 2013; Chen 2012; World Bank 2008	Informal sector (enterprises only) Focus: informal enterprises' contribution to formal GDP	All activities by unincorporated household enterprises that do not conform to the exiting legal and administrative framework (legal definition of informality), and that are below a certain number of employees (production definition of informality).	Includes: **underground production, illegal production** (self-employed, family units, and micro-entrepreneurs) Excludes: **household own final use, agricultural activities**
ILO 2013a; ILO 2013b; Eurostat 2013; Chen 2012	Informal economy (labor only) Focus: informal labor as a share of total labor	All employment relationships inside and outside of informal enterprises not legally regulated or socially protected (without secure contracts, worker benefits, etc.)	**Informal employment in informal enterprises** (employers, employees, own-account operators, and unpaid family workers in informal enterprises) **Informal employment outside informal enterprises** (domestic workers, casual or day laborers, temporary or part-time workers industrial out-workers and unregistered or undeclared workers)
***See Appendix A1 for full version of this table, including measures, methods and critiques of each conceptualization.**			

The distinction between informal and illegal activities is an important one made early in the development of informal taxonomies. Feige (1990), classified various underground activity according to the particular institutional rules violated, while de Soto (1989) separated informal and illegal to avoid the damaging conflation of activities carried out by certain criminal (or terrorist) organizations and the activities of the majority "other," which while considered "extra-legal" by tax or regulatory standards, are not criminal in nature.[5]

In the rest of this section we will review the literature on various understandings of informality, as confusing and contradictory as they can be, both to elaborate these differences and to build a foundation for understanding our own conceptualization and use of data since a primary goal of this project is to model the dynamics of informality with respect to the size and impact of (1) informal economic activity and (2) the informal labor force, we first turn to conceptual efforts that are generally connected to empirical ones, regardless of the terminology they use for the informal sector.

Characterizing informal economic activity

The non-observed economy (UNECE 2008; Gyomai 2012)

The non-observed economy (NOE) framework used by the OECD and others grew out of the need for the exhaustive measurement of productive activities in tabulating national accounts. The NOE is actually an umbrella term for five problem areas or types of non-observed activities: underground production, illegal production, informal sector production, production of households for own-final use, and statistical underground (see Box 2.1 for definitions). While exhaustive of the non-observed economy, the categories found underneath its umbrella are not mutually exclusive in a conceptual sense, and their exact boundaries are sometimes difficult to draw. However, when the definitions are applied as an accounting framework, steps are taken to reduce any double counting that might occur.

---

[5] de Soto (1990: 12) does not attempt a particularly succinct definition on the informal economy. Instead he offers a the following characterization:

> *The concept of informality used in this book is based on empirical observation of the phenomenon itself. Individuals are not informal; their actions and activities are. Nor do those who operate informally comprise a precise or static sector of society; they live within a grey area which has a long frontier with the legal world and in which individuals take refuge when the cost of obeying the law outweighs the benefit. Only rarely does informality mean breaking all the laws; most individuals disobey specific legal provisions in a way that shall be described later. There are activities for which the state has created an exceptional legal system through which informals can pursue their activities, although without necessarily acquiring a legal status equivalent to that of people who enjoy the protection and benefits of the entire Peruvian legal system; these are also informal activities.* (12)

Box 2.1: Definitions and classification of the non-observed economy

According to the Handbook [(OECD 2002)], the umbrella term for the non-observed economy covers five major areas:

1. ***Underground production****: activities that are productive and legal but are deliberately concealed from public authorities to avoid payment of taxes or compliance with regulations.*
2. ***Illegal production****: productive activities that generate goods and services forbidden by law or that are unlawful when carried out by unauthorised procedures.*
3. ***Informal sector production****: productive activities conducted by unincorporated enterprises in the household sector or other units that are unregistered and/or less than a specified size in terms of employment, and that have some market production.*
4. ***Production of households for own-final use****: productive activities that result in goods or services consumed or capitalised by the households that produced them.*
5. ***Statistical underground****: defined as all productive activities that should be accounted for in basic data collection programmes but are missed due to deficiencies in the statistical system.*

Source: György and van de Ven (2014)

The purpose of these definitions is to provide a conceptual framework that enables a better representation of a country's economy in their national accounts. Many countries have thus begun to adjust their national accounts in order to reflect estimates of their NOEs (see Appendix A3 for a list of definitions used by individual countries for measuring the NOE).

This framework can also be applied across sectors or industries, offering even greater insights into the variety of informality across countries. However, as we will see in the next section, the results (namely the data we need) provided by participating countries that apply the framework have relatively little country coverage, and without expansion will have limited utility in larger cross-country analysis (György and van de Ven 2014).

Nonetheless, one of the most important aspects of the work by the OECD is that they have developed conceptualization and measurement standards that are not only being used by OECD countries, but by an increasingly large number of other countries including Peru. This is important because it means that official GDP statistics actually include the informal economy for these countries. Whereas at the beginning of this project's work we anticipated that official statistics excluded the informal economy, and we believed that we would be adding estimates of the informal sector to official values to obtain a truer estimate of total GDP, we instead came to realize that we needed to keep the official GDP as the total value and instead represent the informal share of that total.

The Shadow Economy (Schneider and Enste 2000)

In developing their framework for measuring the shadow economy, Schneider and Enste (2000: 78) began with a 'common' definition of the informal economy that defines it as "all economic activities that contribute to the officially calculated (or observed) gross national product but are currently unregistered." They found, however, that this definition lacked the specificity necessary to ask more directed questions about the causes and consequences of the informal economy. Additionally, it conflated both legal and illegal activities, a distinction which since de Soto (1989) has been understood as of great importance. Therefore, to more sharply define the informal economy, Schneider (2012: 6) turned to the factors that motivate actors to operate outside the formal sector. This sharper definition includes:

> *…all market-based legal production of goods and services that are deliberately concealed from public authorities for the following reasons:*
> 1. *to avoid payment of income, value added or other taxes,*
> 2. *to avoid payment of social security contributions,*
> 3. *to avoid having to meet certain legal labor market standards, such as minimum wages, maximum working hours, safety standards, etc., and*
> 4. *to avoid complying with certain administrative obligations, such as completing statistical questionnaires or other administrative forms.*

This working definition shifts the concept of informality from one that is an economic residual, to one that is rooted in the intentions of actors. Now non-compliance with tax laws, mandatory contributions, or government regulations is the litmus for informality. It follows that this interpretation would also exclude activities that are kept underground because of their criminal nature.

The shadow economy is clearly a narrower concept than the NOE, and most closely relates to the NEO's underground production category rather than the umbrella as a whole. However, this raises some significant issues when we look at actual estimates of GDP share from the two sources and see much larger values from Schneider's shadow economy approach; we will return to this below.

Both Schneider's definition of the shadow economy and the NOE's definition of underground production imply that the act of non-compliance is the outcome of weighing the advantages of participating in the formal economy (such as access to public services and benefits) against the disadvantages (such as paying taxes or complying with government regulations). However, many participants in the informal sector, particularly in developing countries, conduct their operations there out of survival. This "survivalist" informal sector (Section 3 discusses more) lacks the basic level of productivity required to find employment in the formal sector. In the case of these participants, non-compliance with taxes and regulations is a secondary symptom of a deeper developmental issue. We will explore this distinction through the definitions and classifications of informal employment below.

Characterizing informal employment (ILO and WIEGO)

Since the definitions of the informal economy focus heavily on the intentions of those engaged with it, it is important to have a clear and comprehensive idea as to who the informal laborers are. A common characteristic, if not defining attribute, of informal labor (and the informal economy in general) is that they suffer from low levels of productivity (Oviedo et al. 2009). As a result, the jobs they occupy are, on average, lower paid than those of the average worker in the formal sector, although that is not to say that the payments they receive are less than what the same job would fetch in the formal sector. On the contrary, in some cases low-skill jobs in the informal sector can fetch higher earnings than would the same jobs in the formal sector where higher-skilled jobs are concentrated. The economic advantage that low-skilled workers can potentially gain by working informally is further increased by not paying taxes (Perry et al. 2007).

Jobs in the informal sector are also characterized, and by some conceptualizations defined by a lack of benefits and protections. Employers in the informal sector are not obliged to provide health care options, nor are they bound by labor regulations. As a result, informal employees may be more vulnerable to shocks arising from health issues and job security (Oviedo et al. 2009).

Early definitions of the informal labor force were tied directly to previous formulations of the OECD's non-observed economy framework mentioned above, whereby informal labor was essentially the labor that worked for informal enterprises. In 1993, the International Conference of Labour Statisticians recognized that this enterprise-based definition failed to capture the vast heterogeneity of the informal labor force (ILO 2002). Since then, the International Labor Organization (ILO) and the Women in Informal Employment: Globalizing and Organizing (WIEGO) policy network have led the field in informal labor research. Central to their work is the collection and analysis of household surveys aimed at capturing the multidimensional nature of employment relations globally. Their statistical definition of informal employment provides the framework for the creation of data that quantifies informal employment by gender, sector, and employment relation, as it is found across industrialized, transition, and developing countries. Chen (2012: 7-8) summarizes this definition:

> *For purposes of analysis and policymaking it is useful to, first, sub-divide informal employment into self-employment and wage employment, and then within these broad categories, into more homogeneous sub-categories according to status in employment, as follows:*
>
> ***Informal self-employment…***
> - *employers in informal enterprises*
> - *own account workers in informal enterprises*
> - *contributing family workers (in informal and formal enterprises)*
> - *members of informal producers' cooperatives (where these exist)*

***Informal wage employment**: employees hired without social protection contributions by formal or informal enterprises or as paid domestic workers by households…*

- *employees of informal enterprises*
- *casual or day labourers*
- *temporary or part-time workers*
- *paid domestic workers*
- *contract workers*
- *unregistered or undeclared workers*
- *industrial outworkers (also called homeworkers)*

Furthermore, job status can be classified as operating within the formal, informal, or households sectors, from which the ILO draws the simplified groupings of (1) informal work residing inside the informal sector, and (2) informal work residing outside the informal sector. Box 2.2 further elaborates on this distinction.

Box 2.2: Sectoral distinctions of informal labor

Persons employed within the informal sector (including those rare persons who are formally employed in the informal sector):

- *own-account (self-employed) workers in their own informal enterprises*
- *employers in informal enterprises*
- *employees of informal enterprises*
- *contributing family workers working in informal enterprises*
- *members of informal producers' cooperatives*

Persons in informal employment outside the informal sector:

- *employees in formal enterprises not covered by national labour legislation, social protection or entitlement to certain employment benefits such as paid annual or sick leave*
- *paid domestic workers not covered by national labour legislation, social protection or entitlement to certain employment benefits such as paid annual or sick leave*
- *contributing family workers working in formal enterprises*

Source: http://wiego.org/informal-economy/concepts-definitions-methods

Figure 2.1 represents these distinctions with a matrix in which purple checkers squares indicate formal jobs, light blue indicate informal jobs, and dark grey squares indicate impossible job types. Cells 1 through 6, and 8 through 10 represent informal employment, which includes the jobs represented by square 7 even though the job itself may be considered formal. Cells 3 through 8 represent informal employment residing inside the informal sector, whereas cells 1, 2, 9, and 10 represent informal employment residing outside the informal sector.

Production units by type	Jobs by status in employment								
	Own-account workers		Employers		Contributing family workers	Employees		Members of producers' co-operatives	
	Informal	Formal	Informal	Formal	Informal	Informal	Formal	Informal	Formal
Formal sector enterprises					1	2			
Informal sector enterprises*	3		4		5	6	7	8	
Households**	9					10			

Figure 2.1: Components of the Informal Sector and of Informal Employment. *Note: (*) As defined by the 15th International Conference of Labour Statisticians (excluding households employing paid domestic workers). (**) Households producing goods exclusively for their own final use and households employing paid domestic worker*.
Source: *Jütting and de Laiglesia (2009: 29)*

This employment-centric definition has given new perspective to the variable causes and consequences of informal employment, and has led to new understandings of the dynamics and composition of informal employment. One example is that by shifting the frame of reference to employment relationships the pool of informal job types was found to be more diverse than previously expected. Earlier conceptions of informal labor saw the work as primarily survival activities; now it is seen as including not just these activities, but also multiple forms of wage employment across both sectors of the economy (Chen 2007). This insight also suggests the existence of a hierarchy of earnings and poverty risk in the informal sector that is dependent by sex and type of work (summarized in Figure 2.2 below [Chen 2012]). Distinctions like these become more relevant in our discussion surrounding survivalist and entrepreneurial drivers of informality in Section 4.

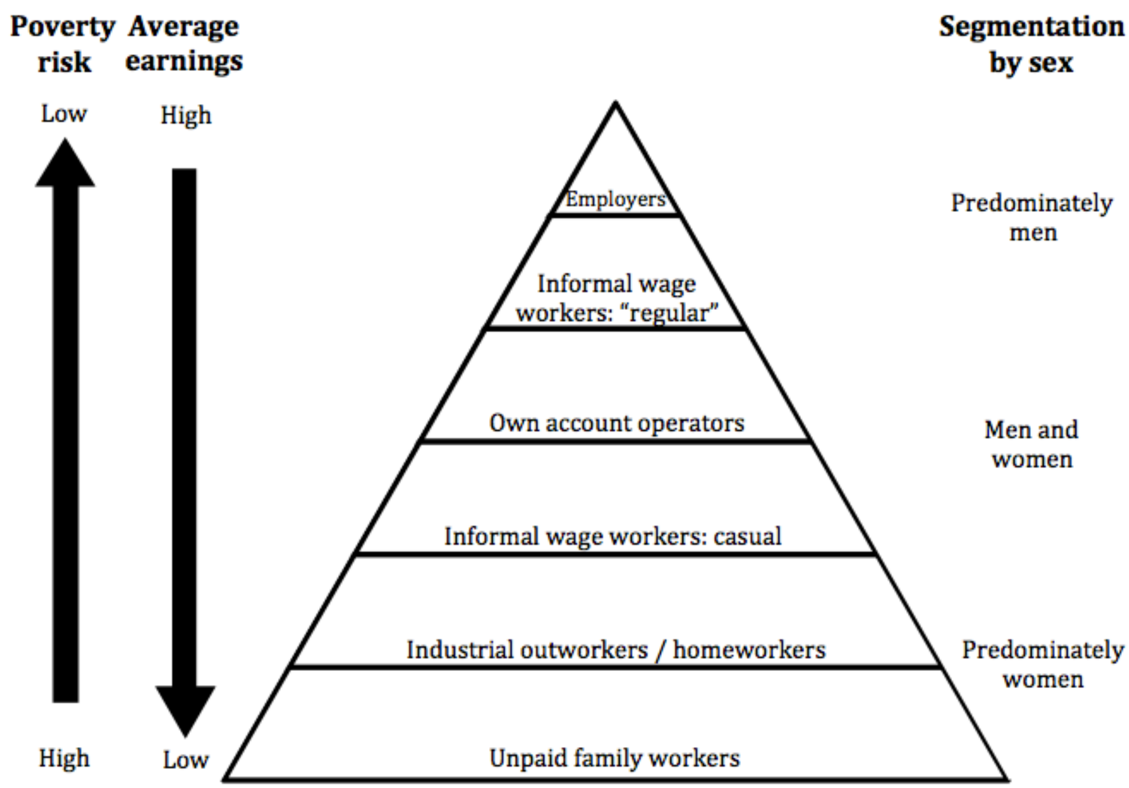

Figure 2.2. WIEGO model of informal employment: Hierarchy of earnings and poverty risk by employment status and sex
Note: Unpaid family workers does not include activity such as that of wives or husbands maintaining the household; it refers to work by family members for the informal market economy outside the household.[6]
Source: Adapted from Chen (2012: 9)

Our own conceptualization

How do *we* define the informal economy? The confusing, and sometimes-contradictory nature of the various conceptions and definitions described in this section certainly point to the complexity of the issue. But their commonalities also provide an important way forward for constructing our own conceptualization of the informal economy. We thus conceive of the informal economy as being comprised of two parts, informal production *and* informal labor. For informal production, we include all informal production \ both within and outside the informal sector. The OECD's non-observed economy framework (NOE)—official adjustments to national accounts—provides the best method for measuring informal production..[7] While this measure does sometimes include illegal production, household production for own-final use, and other activities not reported due to statistical errors or

[6] See ILO (2013a: 24). Similarly, Schneider and Enste (2000: 79) claim not to include contributions from "unpaid or 'pure' household production."
[7] Section 4 returns to the issue of the implications of using NOE data and a conceptualization excluding illegal activities.

deficiencies, their contribution to the overall NOE is small compared to underground and informal production, which closely align with the Shadow Economy concept used by Schneider and others.[8] For informal labor, we follow the lead of the ILO and WIEGO to include not just informal labor employed within informal enterprises or self-employed, but also informal labor in the formal sector and (paid labor) in households (see Box 2.2 for a breakdown of the informal labor categories used in our conceptualization). The NOE framework assumes that informal labor in the formal sector is fully accounted for by standard NSA reporting practices, but we feel this assumption obscures an important aspect of the overall informal economy

The informal economy and its participants strive to remain hidden, a fact that makes measuring their complex and multidimensional nature challenging. For this reason, different definitions of the informal economy have evolved in part through adaptation to insights provided by the very data they helped create and the questions they posed. This iterative process of conceptual and data interaction continues, and, with new datasets and estimation techniques at the disposal of researchers, it appears to be progressing more rapidly. In the next section we will explore the data and assess their utility in supporting a model of the informal economy based on our conceptualization. Then, once we have moved from concept to specification, we can go on to initialize our model and study both the drivers and forward impacts of informality.

Box 2.3 The shadow economy as an alternative measure to the informal economy

While we believe that our approach, based on the NOE framework and survey data, is best for modeling and forecasting the informal economy, we have also added the capability to the International Futures model to forecast the share of the shadow economy as well as the NOE-based informal economy. The share of the shadow economy is initialized with estimations developed by Elgin and Oztunali (2012)—a 161-country dataset on the GDP share of the shadow economy covering the years 1950—2009; for those countries lacking values we estimated their share based on the relationship between the shadow economy size and GDP per capita (more detail on the estimates and the shadow economy more generally can be found in this and the following sections). The process for initializing and forecasting the shadow economy share of GDP is similar to that of the informal economy GDP share and is detailed in Section 6. The primary difference is that under the IFs Base Case, only the informal economy GDP share is used to drive the forward linkages from informality in the model. We have therefore added a switch allowing users to tell the model to drive forward linkages with the shadow economy share instead (see

[8] Furthermore, the estimation methods and resulting levels of informal activity reported in INEI (2014) are far more closely aligned with the OECD's NOE accounting framework than with the model-based estimation techniques used by academics such as Schneider and Elgin.

Section 6 and the Scenario manual for more details). Though this report is primarily focused on the non-observed economy, we also include analysis on the shadow economy and provide detailed documentation on its modeling and use.

3. Moving from Conceptualization to Measurement

In the previous section we covered three prominent conceptualizations of the informal economy: the non-observed economy, the shadow economy, and informal employment. In this section, we will explore the different methods and data used in the literature to estimate the size and composition of the informal economy. Unfortunately, the array of different definitions of informality and the inherent difficulty of observing (and accurately recording) informal economic processes have resulted in a multitude of estimation approaches. This section will introduce what we consider to be the primary methods and datasets used to measure informality and will briefly discuss some of the strengths and weaknesses of each before summarizing the data upon which we primarily rely for our own conceptualization and forecasting.

Estimations of informality

The informal economy is evasive by definition, making its observation and quantification incredibly challenging. While over the years there have been many methods to measure or estimate its size, three dominant approaches have emerged: (1) direct methods such as voluntary surveys or audits, (2) indirect methods which use various economic indicators as proxies for the informal economy, and (3) model-based approaches such as the Multiple Indicators Multiple Causes (MIMIC) model, which uses a system of equations to fit multiple input variables to multiple indicator variables by way of a latent variable that represents the informal economy.

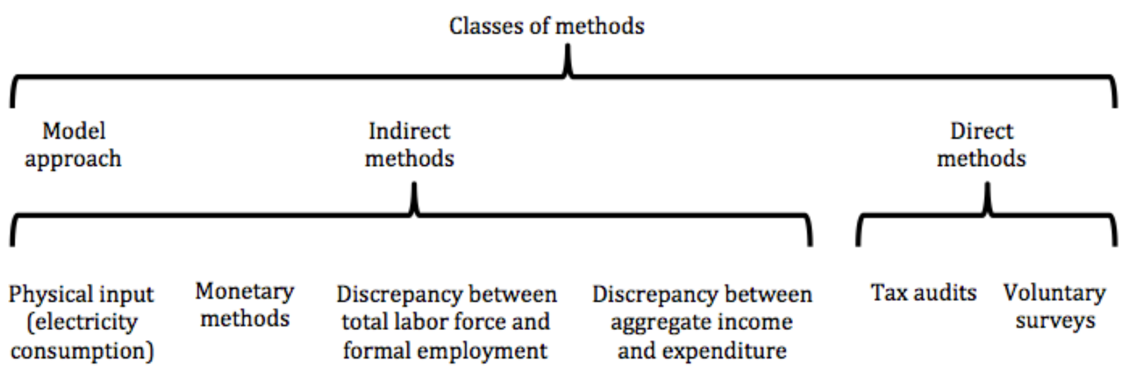

Figure 3.1. Taxonomy of methods to estimate the informal economy
Source: Adapted from García-Verdú (2007) as cited by Perry et al. (2007: 28)

The estimations of different aspects of informality tend to rely differentially on these techniques. As we will see, estimates of the size of informal economy's GDP contribution or value-added frequently rely on more complicated indirect or model-based techniques, whereas measurements of informal employment allow for the use of more straightforward direct or indirect methods (for more detail on the methods of estimation, see Appendix A2).

Box 3.1: Common limitations to data collection / estimation

- ***Limited country coverage:*** *The different estimates of firms' participation in informal activities (e.g. share of annual sales and workers not reported to tax authorities etc.) cover only up to 15 countries. Estimates of informality based on household data (e.g. number of hours in undeclared work etc.) cover most European countries. The proportion of self-employed and multiple job holders in total employment cover the largest set of countries, but they are relatively poor measures of the informal economy.*
- ***Lack of time dimension:*** *Most of the measures of informality are only available at one point in time, with the exception of self employment, which precludes convincingly analysing the drivers and consequences of informality.*
- ***Measurement problems:*** *Proxy measures of informality, such as self-employed, multiple job holders and estimates based on survey data (e.g. share of workers carrying out undeclared work), capture a host of formal activities in addition to informal activities. These measures are also generally incomplete since, at best, they only capture participation in informal activities (and do not account for the intensity of participation) and also suffer from measurement problems (e.g. related to sample stratification, misreporting etc.).*
- ***Consistency across measures:*** *A positive correlation is to be expected between the various measures of informality. While this expectation is borne out in some cases, the cross-country correlation is low, or in some cases even negative, for others… These findings might reflect the tendency for proxy variables to capture different aspects of informality… as well as differences in the participation and intensity of participation in the informal economy across countries.*

Source: Andrews et al. (2011: 22)

Direct "survey" methods

Direct methods are micro-level approaches that use data from household surveys, questionnaires, census data, and tax audits to estimate total economic activity and its formal and informal components (Eurostat 2013). The direct method can provide very detailed data about the composition of the informal labor force, but it is also sensitive to the ways surveys are carried out and the willingness of respondents to provide accurate answers. It also often requires either the expansion of regular surveys to better cover the informal sector or entirely new surveys aimed specifically at the informal sector. Similarly, tax audits can be used to calculate the amount of undeclared income in the economy, but such methods only capture the income disparities from successful audits, and say little about the full extent of hidden or undeclared income that is never caught (Schneider and Enste 2000).

Indirect "indicator" methods

Because informal actors tend to be uncooperative by nature, direct survey methods can be an unreliable data source. Thus, researchers interested in broader estimates of the informal economy have come to rely more often on indirect methods. Indirect methods begin with a particular proxy variable and calculate the difference between what the variable "should be" given no informal activity and what it is actually observed to be. The difference is then assumed to be explainable by the size of the informal economy.[9] Schneider and Este (2000), for example, identify five types of indirect methods: (1) discrepancies between national expenditure and income statistics; (2) discrepancies between official and actual labor force; (3) the transaction approach (discrepancies between official GNP and total nominal GNP); (4) the currency demand approach (discrepancies between currency and tax burden); and (5) the physical input method (discrepancies between official GDP growth and growth of electricity consumption). Eurostat (2013) provides a more condensed list, covering discrepancy methods, monetary methods, and physical input methods. Each method has its own strengths and weaknesses, but all of them also suffer from the fact that they only treat one indicator or driver of informality at a time—attributing to it all changes in the entire informal economy.

Model-based approaches

Model-based methods make use of structural equations to link latent, unobservable variables, such as the size of an economy's informal sector, to observable variables that are assumed to capture its causes and consequences. The Multiple Indicators and Multiple Causes (MIMIC) model has emerged as the preferred model-based technique to estimate the size of the informal economy. The causes used to specify the model generally include observable variables that measure levels of taxation, inflation, salaries, government (including regulation and consumption levels), and unemployment. Indicators, or consequences, generally incorporate the results of one of the previously mentioned indirect methods, most commonly estimation based on currency in circulation. These models also include a reference variable, often real GDP, which is used to normalize the informal economy estimate (Macias 2008).[10] The complete model is a system of simultaneous equations, with one set of

---

[9] For a discussion of these indirect approaches as well as other indirect methods used to estimate the size of the informal economy see Appendix A in Schneider (2005).

[10] Gyomai and van de Ven (2014: 11) indicates that the inclusion of GDP may in fact be problematic, and explain this in the following passage:

> *Usually, the relevant research disregards the fact that published GDP data already contains estimates of the NOE. Consequently, their estimates are presented in percentages of the official GDP as the entirety of the "shadow economy", implying that the official GDP figures do not capture any of the NOE. However, as explained in the main text of this note, official GDP data for most countries, including all OECD countries, already include adjustments for the NOE. Hence, there is an implicit double counting in the results derived from macro-econometric models which can lead to large overestimates of the true size of the NOE.*

Aruoba (2010: 8-9) addresses this potential source of bias in his study regarding the relationship between the informal economy, government policies, and taxes. He explains this treatment, and his findings in the following passage:

> *One issue that needs to be addressed is whether or not official estimates of GDP include any activity that we would label informal. While certain details could differ from country to country, all illegal activities (e.g. drug sales) and most household production activities are excluded from official measures. In our empirical analysis, we make the assumption that the macroeconomic data that we observe reflect only formal activity and do not include any information, either as explicit measurements or as adjustments, about the informal sector. Alternatively, we could make*

equations modeling the effects as a function of the informal economy variable, and the other modeling the informal economy as function of the cause variables (Andrews et al. 2011). The initial output of the model consists of a group of estimated coefficients that are used to create an index. Next, that index is converted into cardinal values necessary for producing a time series of the size of the informal economy (Macias 2008). Given the availability of datasets for these input variables and the proper structure and benchmarking, the MIMIC model is able to estimate the size of the informal economy for a large number of countries over many years. Buehn and Schneider (2012) use this process to create estimates for 162 countries from 1990 to 2007 (more detail regarding the MIMIC model can be found in Appendix A2).

More recently, Elgin and Oztunali (2012) released a panel series for a similar country set, covering 61 years between 1950 and 2010 (see Figures 3.2 and 3.3 for selected countries and global regions). Its creation made use of a two-sector (formal and informal) dynamic general equilibrium model, specified with similar variables to those used in Buehn and Schneider (2012), and calibrated to the 2007 values from the study. A comparison of the two resulting estimates indicates strong overall similarities, and both indicate an average decline in size of informal economies around the world.

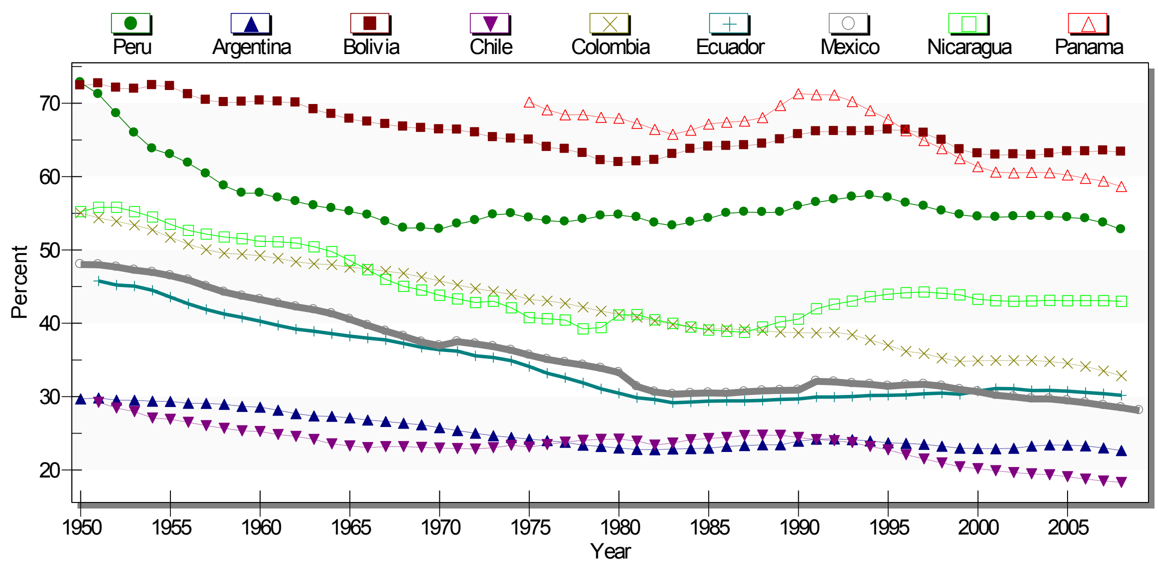

Figure 3.2. Evolution of shadow economy sizes for select Latin American countries according to Elgin and Oztunali (2012), 1950—2008
Source: IFs version 7.14 using data from Elgin and Oztunali (2012)

---

the assumption that statistical agencies fully reflect the level of informal activity in their national account estimates. The resulting ratio of informal to formal activity from this assumption is a simple transformation of the one we use and their correlation is over 0:95: As such all of our quantitative results will be virtually unchanged. In the absence of precise information about how the statistical agency of each country in our sample treat informal activity, we need to make an assumption and while inconsequential, we make the assumption stated above.

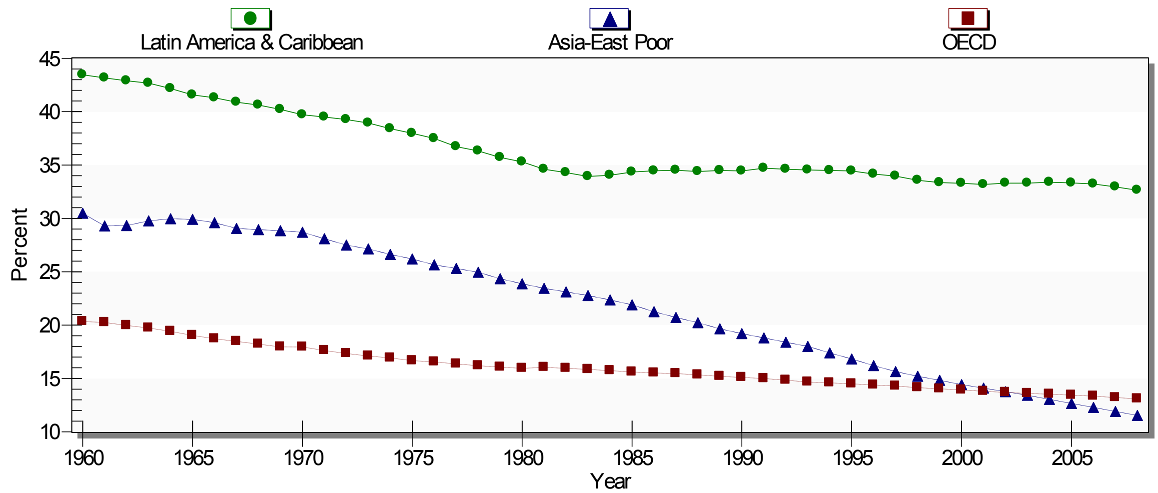

Figure 3.3. Evolution of informal economy sizes for select geographic regions according to Elgin and Oztunali (2012), 1960—2008
Source: IFs version 7.14 using data from Elgin and Oztunali (2012)

The cross-country and time series analysis that this dataset provides has made it very popular among quantitative informal economy research studies (Andrews 2011). Nevertheless it has attracted substantial criticism for its theoretical combination of many assumptions and variables—in particular, Breusch (2005) shows that results are highly sensitive (non-robust) to input variable transformations, concluding that the MIMIC model is "unconvincing as a framework for measuring the underground economy" (Breusch 2005: 28). Dell'Anno and Schneider (2006) address these criticisms, however, and reinforce that "the MIMIC model is still one of the best approaches to this purpose" (Dell'Anno and Schneider 2006: 1). The Buehn and Schneider (2012) estimates and the MIMIC method in general are widely used in quantitative studies of the informal economy.[11]

[11] It is important to point out that since it relies on multiple causal variables to determine the latent informal variable, any subsequent analysis exploring the relationship between these variables and the resulting series might at least partially reflect artificial correlation.

The non-observed economy (NOE) method and results

While the framework provides detailed guidelines for defining and measuring the components of the NOE, different countries still use different combinations of direct and indirect methods (see above) to estimate the extent of non-observed economic activities within their borders (UNECE 2008). Once the level of non-observed activities has been estimated, countries differ again in their use of methods to estimate the GDP generated by the non-observed activities. The UNECE (2008) report identifies the most common approaches: the labor input method, balancing input-output and supply-use tables, and discrepancy methods.

GDP estimating methods in detail:

(1) Labor input method:

> *Labour supply and demand data are compared to estimate inconsistencies in recorded labour… A surplus of labour supply from household surveys over labour demand from business surveys is an indication of non-observed production. Estimates of output and value added per unit of labor input are calculated and then applied to data on labor input to account for output by unregistered and hidden labour. (*UNECE 2008: 6)

(2) Input-output and supply-use methods: *rely on data on the supply of inputs—raw materials, labor, land, fixed capital stock—and use input/output and input/value-added ratios to calculate value-added generated from the inputs* (Eurostat 2013)

(3) Discrepancy methods: Calculates discrepancies in theoretical value added tax (VAT) and actual VAT, theoretical income tax and actual income tax, etc.

Because of the differences in methods between countries, organizations like the OECD use surveys to gather each country's estimates of NOE size, GDP contribution *and* the datasets and methods used by each (UNECE 2008; Gyomai 2012). Two main surveys have compiled non-observed economy (NOE) estimations, in which participating countries were asked to provide breakdowns of each of the five areas of the NOE (recall Box 2.1). To facilitate and harmonize the systematic NOE breakdown across all countries, participants were asked to report estimations according to the Eurostat Tabular Approach to Exhaustiveness. The Tabular Framework is represented in Figure 3.4. Cell colors represent the different NOE areas, where blue is underground (the definition that most closely aligns with Schneider's shadow economy definition), green is informal sector production, red is illegal production, and grey is the statistical underground.

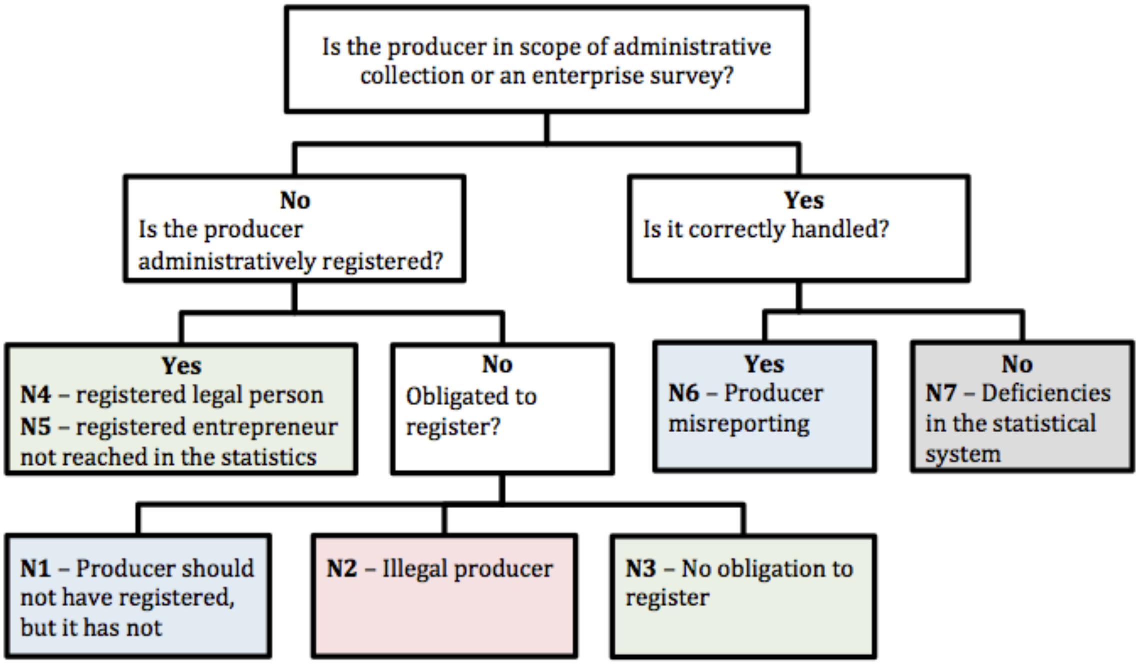

Figure 3.4. The Eurostat tabular approach to exhaustiveness
Source: Gyomai and van de Ven (2014: 2)

Overall, the surveys showed a wide range of GDP share from the NOE in countries participating in the framework. In the UNECE 2005—2006 survey (data years 1995—2003), GDP shares ranged from a low of 0.8% in the US to a high of 31.6% of GDP in Moldova. OECD countries saw the smallest shares, on average, ranging from 0.8% to 15.7% in Poland with an average of 6.6%, while CIS countries saw the largest shares, from 10.7% of GDP in Belarus to 31.6% in Moldova, with an average of 22.7%. Since this dataset reflects the actual adjustments to a country's GDP, we feel that it contains the most defensible estimates of the informal or non-observed economy for policy relevant modeling efforts. Nevertheless there exist some deficiencies that must be considered when working with the data.

As described above, there was considerable variation in how countries handled the allocation of production across categories, as well as the methodologies used to estimate these values.[12] The result is that the NOE GDP share for a given country can vary significantly depending on the methods used—the Czech Republic, for example, saw its NOE share vary from 4.6% under the expenditure approach to 9.3% under

[12] Andrews et al. (2011: 12), explains this further by saying:

> *A caveat is the potential non-comparability as different countries use different methods to estimate the non-observed economy and the estimates differ in their inclusiveness of non-observed activities… Most countries made adjustments for producers deliberately not registering, producers not required to register, misreporting and other statistical discrepancies… Some countries did not state their estimate of the size of the non-observed economy… even though they shared the methods for estimating this sector, perhaps suggesting that the estimates are surrounded by a high degree of uncertainty.*

the output approach. In addition to the cross-country variation in reporting, the data suffer from overall coverage issues. Out of the 35 countries to receive the survey, 19 countries responded with NOE adjustment breakdowns, though some required full or partial confidentiality, and thus were not reported in the survey summary (Gyomai 2012). Moreover, the overall "non-timely nature and the sparse availability only allows cross-sectional, historical macro-economic (and structural) analysis… and is not particularly suitable for identifying more detailed developments over time, or performing analysis in a timely fashion" (Gyomai and van de Ven 2014: 8). To deal with the issues of country and temporal coverage we have blended the NOE series with data from Buehn and Schneider (2012), Elgin and Oztunali (2012), and INEI (2014),[13] resulting in a dataset covering 186 countries, with data available for some or all countries for the years 1960 to 2010 (see Box 3.2). Figure 3.5 shows estimates from three of these sources as well as the resulting blended series for the country of Peru.

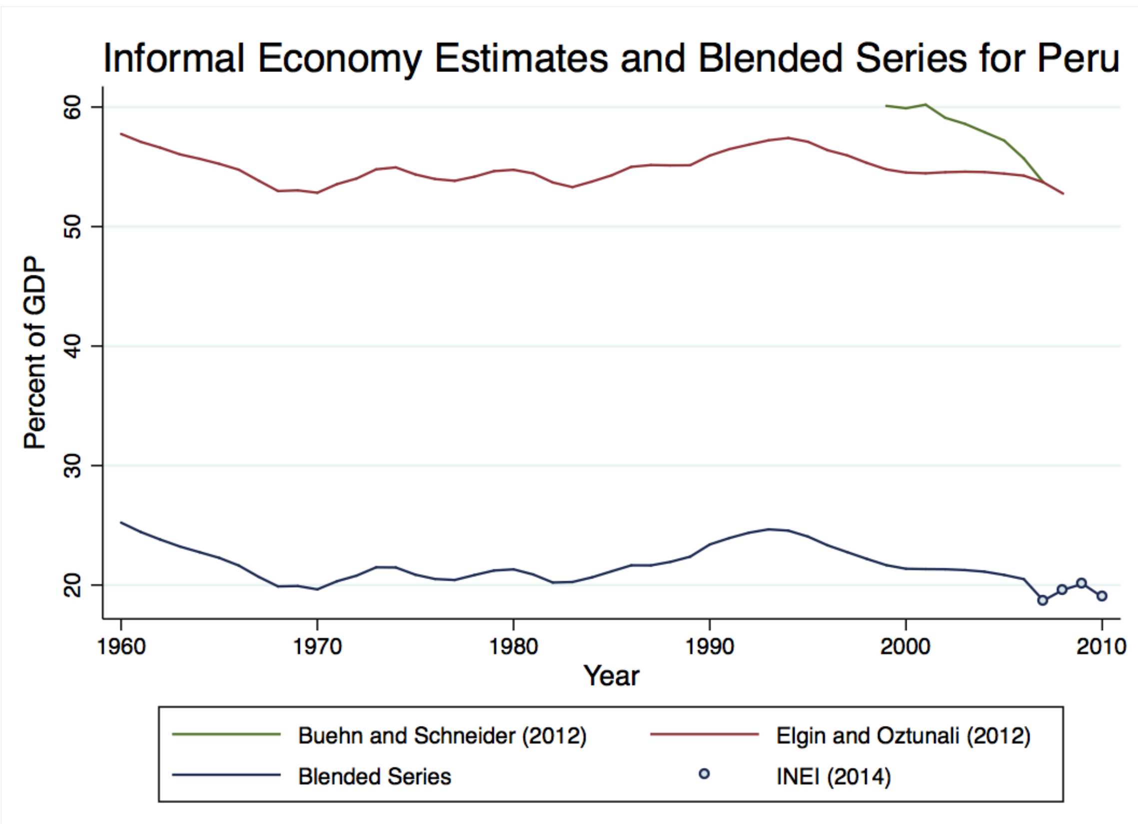

Figure 3.5. Informal economy estimates by individual study and blended series, 1960—2010
Note: The blended series in the figure was created from the three studies cited below, and is used as the primary informal economy series for this project.

[13] For the INEI data, the informal economy GDP share is reported in terms of the Value Added by the informal sector, a subdivision of the overall informal economy as measured by the NOE (see Table 2.1 and discussion on NOE). The informal sector makes up the majority of the NOE.

Source: Buehn and Schneider (2012), Elgin and Oztunali (2012), INEI (2014), and IFs version 7.14

Box 3.2. Conceptual justification and process for blending OECD, INEI, Schneider, and Elgin datasets

There are many approaches to estimating the size of the informal economy and none of them are without flaws. We examined each approach in search of the best dataset(s) to use in our own modeling efforts. The Multiple Indicators and Multiple Causes (MIMC) approach used by Schneider, Elgin, and others, is the most widely embraced macro-modeling method in the literature. The MIMIC approach, however, has come under significant criticism due to its (1) being too simple and aggregate in nature to capture the complexity of systems of national accounts; (2) being very sensitive to initial assumptions, particularly around the use of currency demand methods to re-scale the "latent" size of the shadow economy output so it can be directly compared with GDP; and (3) often resulting in double counting as it assumes official GDP figures do not capture any of the shadow economy even though many countries already adjust their official GDP figures to take the non-observed economy (NOE) into account (OECD 2002; Breusch 2005; Goymai and van de Ven (2014).

The OECD (2002: 187) described MIMIC methods as tending "to produce spectacularly high measures, which attract much attention from politicians and newspapers." For example, they cite results that calculated the size of the shadow economy in Canada for 1990—1992 to be between 10 and 13.5 percent of the country's GDP while an official Statistics Canada report estimated an upper limit of 2.7 percent for the underground economy in 1992. Similarly, Goymai and van de Ven (2014: 11) found that Schneider's estimates of the shadow economy tended, on average, to be 3.6 times larger than NOE estimates by systems of national accounts and 6.7 times larger than the underground economy portion of the NOE (the underground economy is the portion that most closely matches the definition of the shadow economy, and while the underground economy accounts for the vast majority of the NOE, the NOE as a whole captures a larger array of activities than the shadow economy and so ought to be larger than the shadow economy). Goymai and van de Ven (2014: 11) concluded that such great differences were likely due "in great part, [to] unrealistic model assumptions and calibration decisions."

The NOE, while still suffering from some shortcomings, including variation in methodologies across countries, and inclusion of non-reported activities not normally considered part of the informal economy like illegal operations—though such activities make up only a few percent of the overall NOE—is based on detailed survey data tied to national accounts and is therefore more policy relevant than the results produced by macro-models. We also found that both NOE and shadow economy data have a similar (parallel) relationship with per capita GDP, just with the shadow economy being significantly larger.

Given the above justifications, the NOE methodology yields what we feel are the most defensible estimates of the informal economy for policy-relevant modeling efforts. The NOE dataset, however, was lacking when it came to the extent of country and temporal coverage. We therefore blended the NOE data (UNECE 2008) with data from Buehn and Schneider (2012), Elgin and Oztunali (2012) and INEI (2014) to create a dataset that covers 186 countries with data available for some or all countries for the 1960 to 2010 period. Creating this blended data set required us to adjust the MIMIC-derived estimates to bring them in-line with the size of the NOE estimates. To do this, we first plotted regression lines for each data series against the log of GDP per capita at purchasing power parity in USD 2011 (logged so that the relationship is linear) and then did the same for *GDPInformal%Adjustment* (a data series reported by countries using the NOE framework to adjust their GDPs to account for the non-observed economy in their national accounts). We then took the differences between the regression lines and each level of logged *GDP2011PCPPP*, subtracted the differences from each, chained the results for countries that existed in both data sets to available data from UNECE (2008) and INEI (2014), and finally, set a hard limit to ensure that the adjusted values do not fall beneath 1%.

The full blended series gives preference first to the UNECE and INEI series, then the Elgin and Oztunali series, and then the Buehn and Schneider series.

One additional caveat is that the NOE also adjusts for economic activity not reported due to its illegal nature or due to statistical deficiencies, which are two classes not included in our definition of the informal economy. We have decided to proceed with this measure despite the additional activity it captures for a few reasons: First, the results of the surveys show that, "the underground economy almost always represents the most significant part of the adjustments for non-exhaustiveness, reaching as much as 80% of all adjustments in some countries" with the informal economy making up much of the remaining 20% (OECD 2014: 6). Second, using only data from countries that have provided NOE estimates disaggregated by activity class would further reduce the number of countries covered in our dataset. Third, since there is debate as to whether NOE adjustments dramatically underestimate the informal economy (see Schneider and Williams 2013), and since model-based estimates are on average significantly larger, the impact of these additional activities on modeling efforts is likely minimal.[14]

[14] We acknowledge that with this approach we, to some extent, contradict a definition of informal activity that excludes illegal activity. However, Table 3.1 indicates that illegal activity accounts for on average less than 3% of total NOE adjustments, or less than a one third standard error. Furthermore, disentangling the illegal from other underground activities can prove difficult, as their borders are often times unclear particularly when comparing across countries (INEI 2014).

	Underground N1+N6	Illegal N2	Informal sector N3+N4+N5	Statistical deficiencies N7	Total NOE
Austria	2.4 (31.7%)	0.2 (2.1%)	1.5 (19.4%)	3.5 (46.8%)	7.5 (100%)
Belgium	3.8 (83.8%)			0.7 (16.2%)	4.6 (100%)
Canada	1.9 (88.2%)	0.2 (8.2%)		0.1 (3.6%)	2.2 (100%)
Czech Republic	6.3 (77.6%)	0.4 (4.5%)	1.3 (15.6%)	0.2 (2.3%)	8.1 (100%)
France	3.7 (54.7%)		2.9 (42.7%)	0.2 (2.7%)	6.7 (100%)
Hungary	3.1 (27.9%)	0.8 (7.5%)	3.1 (28.6%)	3.9 (36%)	10.9 (100%)
Israel	2.2 (32.6%)		1.4 (21.8%)	3 (45.6%)	6.6 (100%)
Italy	16.2 (92.8%)			1.2 (7.2%)	17.5 (100%)
Mexico	5.5 (34.7%)		10.4 (65.3%)		15.9 (100%)
Netherlands	0.8 (36.6%)	0.5 (20.1%)	0.5 (20%)	0.5 (23.2%)	2.3 (100%)
Norway	0.5 (51.5%)	0 (0.3%)	0.5 (43.8%)	0 (4.4%)	1 (100%)
Poland	12.7 (82.6%)	0.9 (6%)	0 (0%)	1.8 (11.4%)	15.4 (100%)
Slovak Republic	12.1 (77.3%)	0.5 (3%)	2.9 (18.7%)	0.2 (1%)	15.6 (100%)
Slovenia	3.9 (38.2%)	0.3 (3.2%)	2.8 (27.7%)	3.1 (30.9%)	10.2 (100%)
Sweden	3 (100%)				3 (100%)
United Kingdom	1.5 (65.6%)		0.5 (22.9%)	0.3 (11.4%)	2.3 (100%)

Table 3.1. NOE adjustments by informality-type, percentage of GDP, share of adjustment type within total NOE
Source: Gyomai and van de Ven (2014: 6)

Measuring the size and dimensions of the informal labor force

Direct and indirect methods lend themselves more to the estimation of the more "countable" informal labor force, where household surveys, questionnaires, and census data can measure the level of participation in different types informal positions for a sample population. When direct measurements were unavailable indirect methods used various residual-balancing techniques to estimate the size of informal labor.[15] Collection and calculation efforts by the ILO and partners like WIEGO have led to the production of a new database of informal employment covering 40 countries (ILO-WIEGO henceforth), and differentiating across the following three dimensions of employment relations: (1) informal employment inside the informal sector, including formal employment in the informal sector, (2) informal employment in formal enterprises, and (3) informal employment in households, such as domestic employment (see Figure 3.6). The latter two are often combined, forming a group of informal employment outside the informal sector.

[15] For a description of these techniques see ILO (2013).

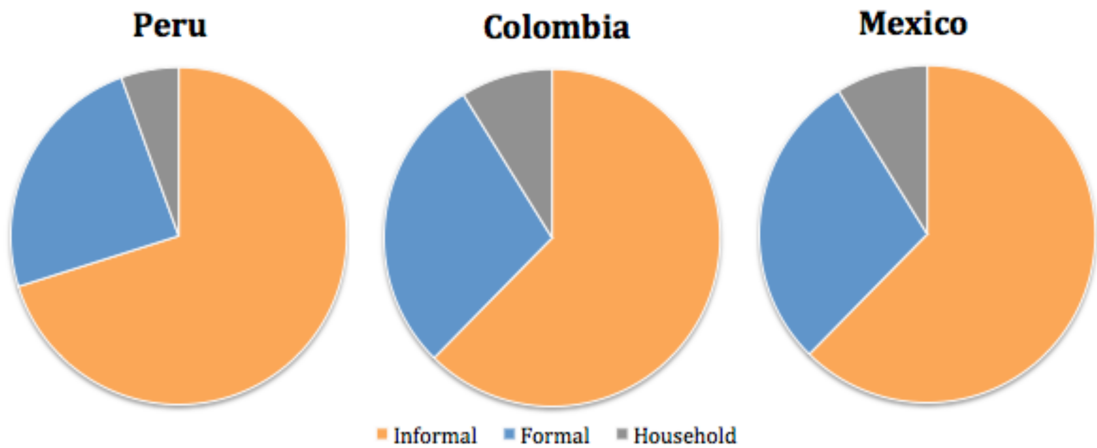

Figure 3.6. Informal employment shares across the informal sector, formal sector, and households for Peru, Colombia, and Mexico
Source: Data from ILO-WEIGO

The World Bank also maintains a historical dataset measuring the sex-disaggregated share of non-agricultural informal employment for a similar set of countries. We combine these two series for a measure of informal labor as a percent of total labor in Figure 3.7 (See Box 3.3).

Box 3.3 Creating a combined historical dataset of informal labor

The World Bank's historical data on informal labor comes from their *Gender Statistics Database* (available at http://data.worldbank.org/data-catalog/gender-statistics). Though the database itself does not include any metadata concerning the source of this data, the 2014 World Development Indicators report cites the ILO's *Key Indicators of the Labour Market* as the original source of the data (36). This suggests that the WB and ILO-WIEGO datasets are based on the same methodology and definitions, making them directly comparable. The World Bank data covers a period from 1990 to 2011 while the ILO-WIEGO data spans two periods, from 2004 to 2012 and from 2009 to 2013 (ILO 2014a and 2014b). To blend these series we began with the most recent ILO-WIEGO data (2009 to 2013), appended the earlier years (2004—2008) and filled in missing values for 2009 to 2012 from the prior ILO-WIEGO dataset, and then did the same for the World Bank data, appending the years 1990—2004 and filling in any missing values from 2004—2011. Because the World Bank data is only presented as disaggregated by gender, we therefore took the intermediary step of computing a combined total series from the male and female percentages.

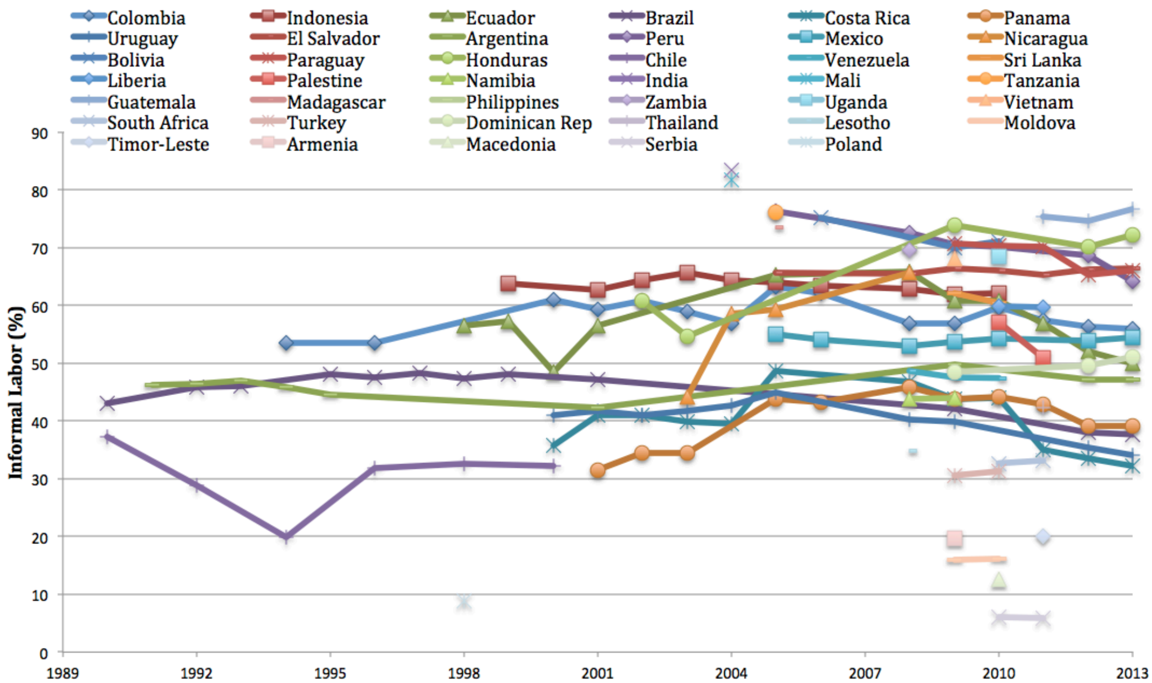

Figure 3.7. Informal labor as a percent of total labor over time, 1989—2013
Source: Data from ILO-WIEGO and World Bank

Furthermore, for around 30 countries, the ILO-WEIGO data is further disaggregated by gender and economic sector, permitting an even more useful snapshot into the heterogeneous nature of informal labor. Figure 3.8 shows the distribution of informality across all non-agricultural activities (in blue), and is also broken down by five economic sectors for Peru. The graph illustrates that informality is extensive in all areas of the Peruvian economy. Figure 3.9 shows the share of persons in informal employment outside the informal sector, disaggregated by sex.

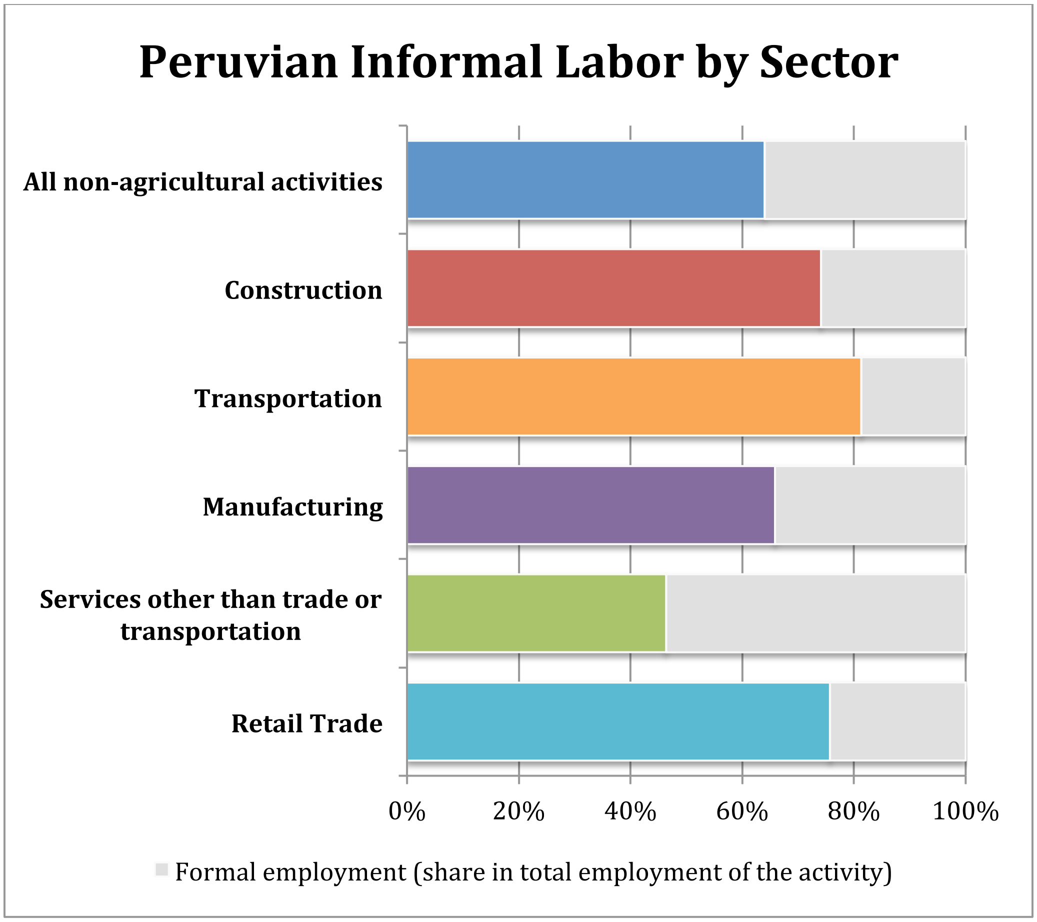

Figure 3.8. Informal share of all non-agricultural labor, and by sector for Peru, 2013
Source: ILO (2012)

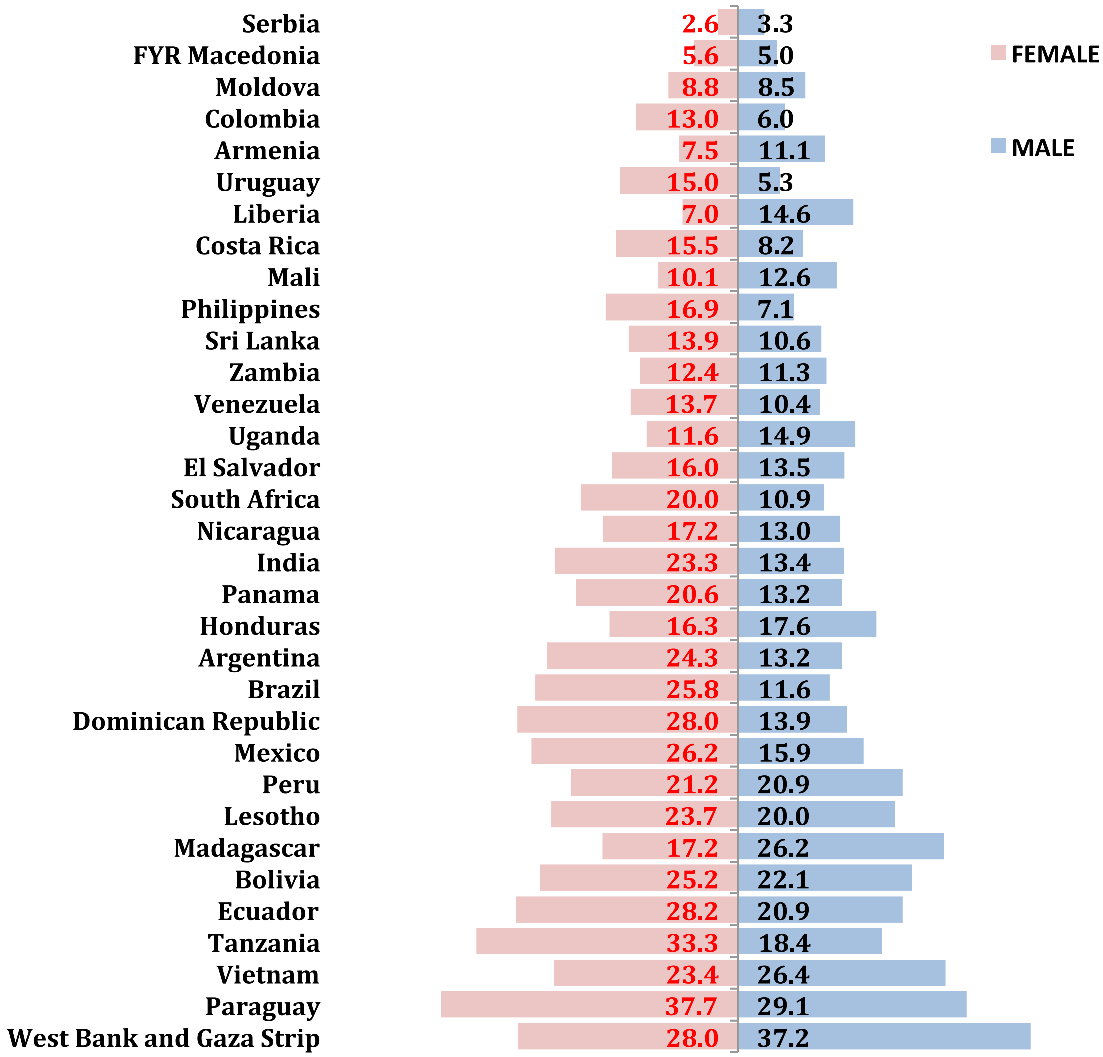

Figure 3.9. Share of persons in informal employment outside the informal sector by country and sex, latest year available
Source: ILO (2012)

These informal labor series, however, also exhibit some deficiencies related to country and temporal coverage. Of the countries found in the ILO-WEIGO study, only 4 are OECD, providing little representation of informality in higher-income states. Also, while the informal labor as a percent of total labor (for men, women, and total) data covers the years 1990 to 2011, most countries have only one or two data points—providing little insight as to whether informal labor is growing or shrinking. The cost and time requirements of developing and administering the surveys necessary for these series also generally prohibit more frequent updates (Gyomai and van de Ven 2014). Nevertheless, the methods used in the collection and production of this dataset offer the strong empirical anchoring that model-based estimates lack.

Comparing measures of the informal economy

Table 3.2 compares the series mentioned in this section for a selected group of Latin American and Caribbean countries. Values represent an average of available data over the years 1990 to 2011. One feature that stands out immediately is that in nearly all cases, informal labor is larger than informal GDP share, supporting the common assessment that the informal economy is less productive than the formal economy.

Country	Informal GDP (% total GDP)		Informal GDP Increment (points)	Informal GDP Increment (points)	Informal Labor (% total labor)
	Schneider	Elgin	NOE Adjust	Blended Series	ILO-WIEGO
Argentina	25.3	25.2		10.7	46.0
Bolivia	66.1	66.2		55.3	72.1
Brazil	39.0	44.7	12.8	13.2	46.4
Chile	19.3	23.4		8.5	30.5
Colombia	37.3	42.6		24.6	58.5
Costa Rica	25.7	32.0		14.6	41.6
DominicanRep	31.9	43.5		23.1	48.5
Ecuador	32.4	34.3		19.5	58.7
El Salvador	45.1	53.4		36.2	65.8
Guatemala	50.5	59.3		41.0	75.4
Honduras	48.3	57.2		40.3	63.1
Mexico	30.0	35.8	12.1	12.9	54.0
Nicaragua	44.6	45.0		34.7	56.9
Panama	63.5	65.8		52.6	40.5
Paraguay	38.8	46.4		25.9	70.5
Peru*	58.0	56.5	(19.35)	21.9	(69.6)
Uruguay	50.6	52.2		36.0	41.7
Venezuela	33.8	30.8		17.9	47.9

*** Values in parenthesis are taken from INEI (2014).**

Table 3.2. Comparisons of different measures of informality for Latin American countries

Note: When time-series data exists, values are averaged from 1990 to 2011 (or for any data found within that range). (*) The values in parentheses for Peru come from INEI (2014: 53).[16] The original INEI value of 15.5 percent includes the informal sector only and not the informal economy as a whole, thus the percentage has been adjusted to be compatible with the wider NOE measure. Thus, Peru's NOE Adjust value (19.35) is for the entire economy. The NOE adjustments are increments of the informal economy to the formal one; therefore they are not strictly comparable to the estimates of informal shares (a 20 percent increment would mean that the share was 20/(100+20)=16.7%).

Source: The datasets include Buehn and Schneider (2012), Elgin and Oztunali (2012), UNECE (2008), ILO (2012) and the core blended informal economy series from IFs version 7.14

---

[16] The INEI values for Peru are considerably different from the values from Schneider and Elgin with respect to GDP share and to a lesser extent with respect to values from ILO-WEIGO on the labor share. The model is now using the values from these other sources and not those from INEI; this is a matter for discussion with our colleagues at CEPLAN. The disadvantage of using INEI numbers would be inconsistency with those used for other countries, however we believe this inconsistency is minimized with the switch to an NOE-based data set. The advantage of using them would be consistency with Peruvian government estimates. Substituting the greater ratio of informal labor to informal GDP share in the INEI numbers (69.6 / 19.35 = 3.60, compared to 72.5/57.2=1.27 in the IFs model currently, would mean that a reduction in informal share would have a greater impact on forecasts of Peruvian productivity and growth.

Another interesting observation is that the NOE-based adjustments to the GDP are significantly lower than model-based GDP share estimates from either Schneider or Elgin.[17] The fact that the NOE conceptualization shown earlier in Box 2.1 is more inclusive than that of Schneider and Elgin makes this especially surprising and also raises concerns. We know that country applications of the NOE framework vary considerably. For instance, some include illegal activities and some do not (see INEI 2014: 40-42 for the treatment of illegality by Peru). UNECE (2008: 10) lists those countries that have accounted for it (using informality type N2) and those that have not. Gyomai (2012) presents a similar table that suggests most of the NOE adjustments are dominated by underground production, which should be quite close to Schneider's conceptualization. Yet this apparent general equivalence of OECD adjustment practice with Schneider conceptualization again suggests that the values they produce should be much more similar than they in fact are.

Figure 3.10 elaborates on this Schneider-Elgin/NOE issue and moves from Latin American countries to global coverage. We can see that the two model-based estimates have a very high correlation with each other and relatively strong correlation with the NOE method-based national account adjustments, but are on average nearly 17% points (or around 4.5 times) higher for reporting countries.[18] This is an important consideration to have in mind when using the model-based estimates for analysis.

[17] Note: Peru was not included in either survey mentioned above, though estimations using the same framework can be found a 2014 working document put forth by the INEI.
[18] The NOE data used in this analysis capture the full adjustments those countries have made, and therefore include activities that are excluded by Buehn and Schneider (2012). The OECD's underground production adjustment values are more similar in definition, however have fewer reporting countries and more cross-country methodological variation.

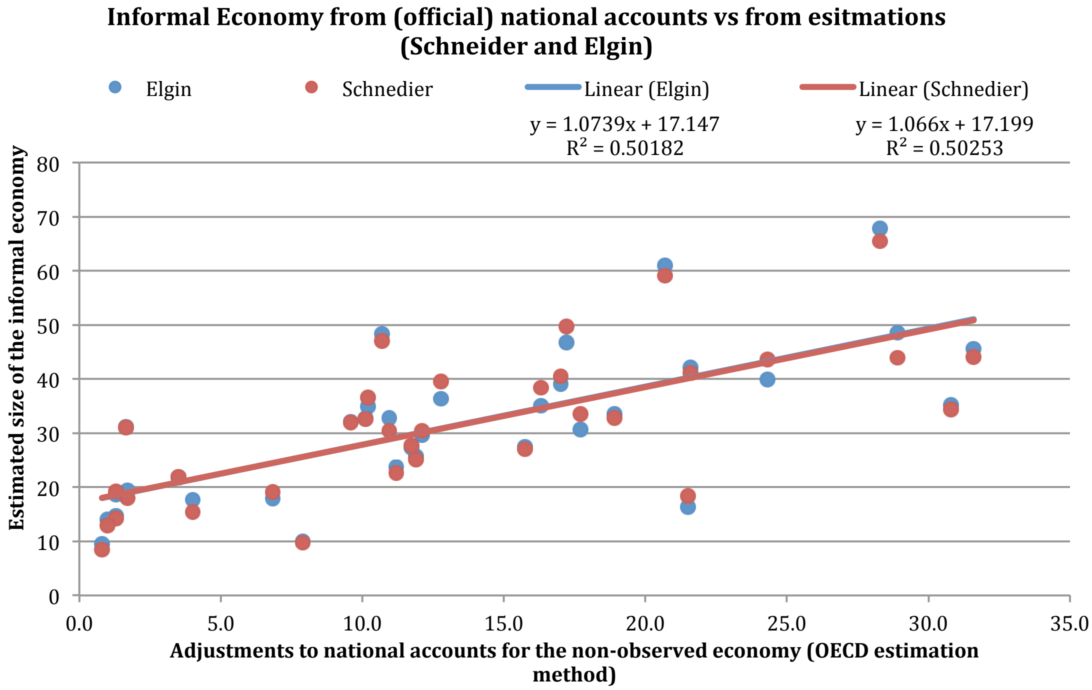

Figure 3.10. Scatter plot illustrating the relationship between Buehn and Schneider's MIMIC estimates, Elgin and Öztunali's dynamic general equilibrium model estimations with the average adjustments to national accounts by reporting countries

Note: The two regression lines are so close to identical that they cannot be distinguished in the graphic.
Source: Buehn and Schneider (2012), Elgin and Oztunali (2012)

In addition to the very different estimates that model-based approaches produce relative to those from applications of the NOE conceptualization, critics have noted that these model-based approaches can show inconsistent or conflicting results when compared against one another.[19] This offers further support for our decision to build our modeling efforts around NOE-based values. Nevertheless, given the shortcomings of model-based estimates, and the relative limited coverage of NOE adjustments, our blended series measuring informal GDP share represents a reasonable effort to frame levels of informal economic activity. Because estimation of informal GDP share is so complicated and different across sources, García-Verdú (2007: 1) argues that "the best measures of the magnitude of the informal sector are the share of informal employment in total employment estimated using household survey data, and estimates of the contribution of the informal sector firms to GDP based on survey data from micro and small firms".

In summary, we use two different primary data series for our analysis and forecasting of the informal sector.[20] For economic value size we rely on the blended

[19] See Gyomai and van de Ven (2014) and Breusch (2005).
[20] In the process of modeling the informal economy in IFs, we developed a capability for importing country-specific values to override those from external data sources used for model initialization across countries. Thus we will be able to replace values from such general sources with ones that experts in countries such as Peru determine to be better estimates.

series of NOE adjustments and model-based estimates—giving preference to levels of informality indicated by the former. For the size the informal labor force we look to the ILO-WEIGO informal labor data series for a strong empirical foundation to represent the informal sector in IFs. In addition to the overall share of the informal labor force, we also use the breakdowns of informal labor by sector (informal, formal, and household), aggregating them into the categories of inside the informal sector and outside the informal sector. Table 3.3 presents some of the descriptive statistics regarding these variables in their most recent year available.

Variable	Obs	Mean	Std. Dev.	Min	Max
Non-Obs Econ	40	14.17	8.97	0.80	31.60
Inf GDP (Schnedier)	161	31.75	12.20	8.10	63.50
Inf GDP (Elgin)	160	31.03	12.06	8.07	63.34
Inf GDP (Blended)	165	19.39	13.22	0.80	56.32
Inf Lab (% Tot Lab)	41	51.02	21.17	5.90	83.51
Inf Lab in Form Sec	33	23.62	12.19	5.18	50.93
Inf Lab in Inf Sec	33	67.10	13.61	37.33	86.15
Inf Lab in HH	33	9.27	8.74	0.00	44.15
Inf Lab Out Inf Sec	33	32.89	13.60	13.85	62.67

Table 3.3. A statistical summary of the most recent measures of informality
Note: the first three variables are used to blend the fourth, but are not used in direct initialization of the model.
Source: The datasets include Buehn and Schneider (2012), Elgin and Oztunali (2012), UNECE (2008), ILO (2012) and the core blended informal economy series from IFs version 7.14

Further, we give some primary position to the informal labor share data produced by ILO-WEIGO in our model formulation. Subsequent sections will explain how we use the different series for estimation and for the foundation of informal labor share and informal GDP share variables in the IFs model.

4. Drivers

What does the literature say about what drives the size of the informal economy?

> *There is evidence that, in part, informality is merely a stage in the development process: the ubiquitous microfirm reflects the unattractive options in the small modern sector and the traditional reliance on family and community. However, other evidence suggests that, in part, informality is a canary in the coal mine – the symptom of poor policies and, more profoundly, a lack of confidence in the state and perhaps in our fellow citizens.* (Perry et al. 2010: 19)

The informal economy is a complex and multifaceted phenomenon, as evident in its multiple definitions, conceptualizations, and estimations. This is also true for its determinants. Over the years, there have been many unsuccessful attempts to define a generalizable set of causes of informality, and their failure was often due to trying to apply a single, universal definition of the situations that give rise to informality. For example, one analytical approach that has proven useful in untangling some of the discrepancies between a set definition and actual observation is in making the distinction between informality in developed and less developed countries. Through this lens, two faces of informality begin to emerge, each with unique qualities and determinants (Gërxhani 2004).

The differences in informality between developed and developing countries begins with the relationship of informality to GDP per capita. Figure 4.1 shows that the size of the informal economy declines quite regularly with higher GDP per capita at purchasing power parity. Notably, this occurs within as well as between the geographical regions color-coded in the figure. Although the discussion below emphasizes the divide between developing and developed countries, as does some of the literature, the reality, as shown in Figure 4.1, appears to be more of a continuum.

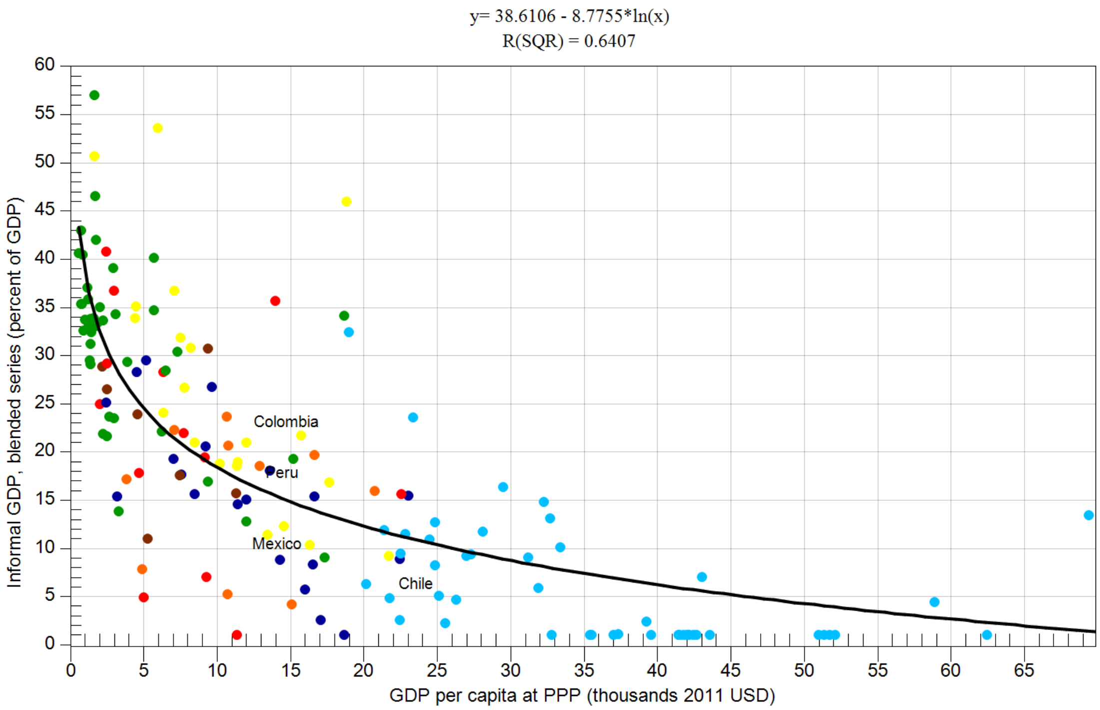

Figure 4.1. Cross-sectional relationship between the blended informal GDP series and GDP per capita at PPP
Note: Both series are using most recent data points. Colors represent different World Bank income groups, where light blue is high-income countries (excluding Qatar), yellow is Latin America and the Caribbean developing countries, green is sub-Saharan African countries, red is East Asia and Pacific developing countries, dark blue is Europe and Central Asia developing countries, orange is Middle East and North African developing countries, and brown is South Asian developing countries.
Source: IFs version 7.14

While an income-based distinction between less- and more-developed societies might appear obvious, given the strong correlation between per capita income and size of the informal sector, it also supports the coexistence of historically conflicting schools of thought: the dualist perspective, which finds greater support in developing countries, sees an unskilled informal sector that operates largely in isolation from the formal world; and the legalist orientation, which has a stronger foundation in developed countries, and is more concerned with explaining informality in terms of the regulatory environment.

Our interest is not just in developing and high-income countries at any given point in time, however, but also very much in the possible transition of countries through the middle-income passage. And while we recognize that the character and dominant drivers of informality in developing and high-income countries may differ significantly, there will also be overlap. Thus, even while our discussion here draws out the insights of those who focus on one grouping or the other, we have actually built a framework that integrates insights across both groups.

Within underdeveloped countries, operating in the informal sector is often a matter of survival. Formal jobs are often very limited and even where they do exist, individuals, and thus households, may be excluded from such employment by failing to have the human capacity required. In contrast, within developed countries, an actor's decision to participate in the informal sector is more likely determined by weighing the costs and benefits of exiting the formal sector. For the sake of simplicity, we will refer to the former situation as "survivalist" and the latter as "entrepreneurial."[21]

Again, it is important to emphasize that these categorizations do not precisely characterize the informal sector of any given country, nor do the two categories exist in isolation or opposition to one another. These varieties of informality can be found in combination across all countries. The distinction is nevertheless useful in providing a typology for discussing the multiple causes of the informal economy and for helping us determine where to focus and how to organize our empirical analysis.

Survivalist economies and the transformative power of deep developmental drivers

Poverty and low levels of productivity often characterize survivalist actors. Their economic activities generally take the form of subsistence production or the sale of basic goods and services. While many of these goods and services may ultimately be consumed by or provided to actors in the formal sector, survivalists often lack the higher-level skills necessary to secure the higher wages, benefits, and security of formal employment. Moreover, those who enjoy more mobility may voluntarily return to the informal sector where wages can actually be more attractive for low-skilled workers.[22] The root of the challenge facing survivalists can be traced back to the most basic developmental obstacle: the inability to accumulate human capital. Figure 4.2 shows the relationship between years of formal education attained by adult populations and the relative size of the informal economy.

[21] See Chen (2012) and Perry et al. (2007) for further discussion on the causal theories around "exit" and "exclusion" from the formal sector.

[22] Docquier et al. (2014) explains this with the following rational:

> *low-skilled workers are mobile across sectors whereas high-skilled individuals only work in the formal sector. When the number of high-skilled workers is small, there is little demand for low-skilled labor in the formal sector and formal firms pay low wages to the less educated. Many low-skilled workers then move to the informal sector where wages are more attractive. Informality thus serves to protect low-skilled workers against very low levels of income offered in the formal sector and extreme poverty.* (5)

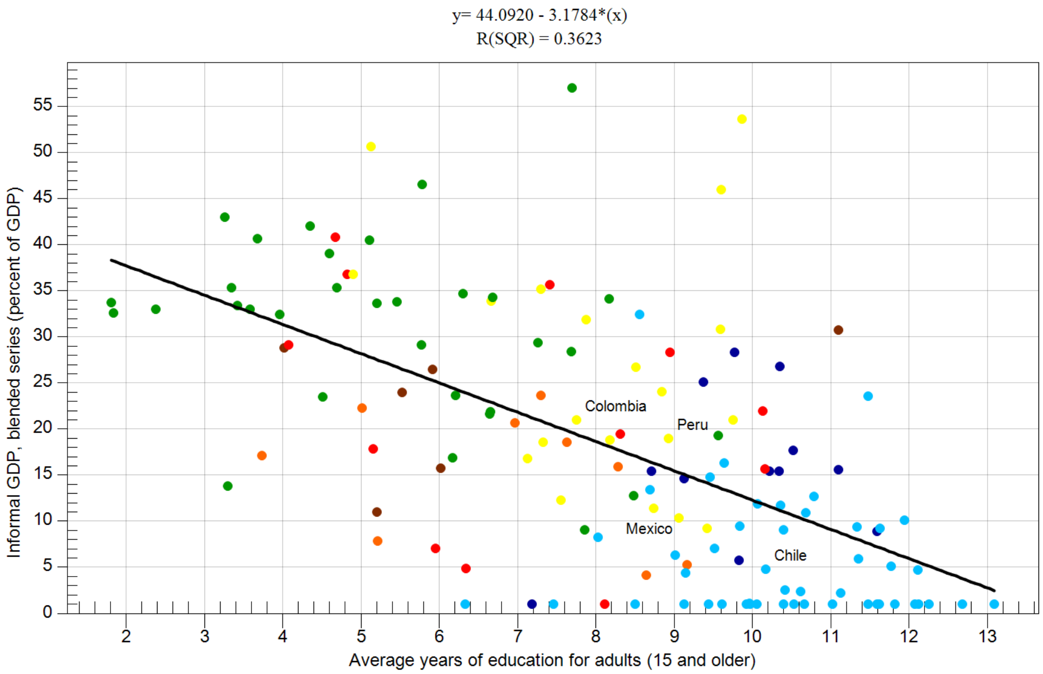

Figure 4.2. Cross-sectional relationship between informal GDP as a percent of total official GDP (blended series) and average years of education for adults 15 and older
Note: Both series are using most recent data points. See Figure 3.1 for explanation of the color-coding.
Source: IFs version 7.14

As we will discuss below, education and GDP per capita are highly correlated and have roughly comparable power to help explain change in informality. Development of human capital such as education is, however, a leverage point for societies as they advance into and through the middle-income transition and even for high-income countries as they seek to maintain their competitive position globally. Growth in GDP per capita, by comparison, is a consequence, not a policy lever, hence, our focus on education as the deep, long-term driver.

Jütting and Laiglesia (2009) and Tansel (2012) show that human capital, and particularly educational attainment, are major determinants of informal/formal labor mobility. With the increasing returns to labor productivity that already exist at lower levels of education, and the promise of higher wages and benefits in the formal sector, the incentives for households to invest in education appear obvious.

Households in informal economies, however, face a number of obstacles, on both the cost and the benefit side. Informal actors in developing countries not only often reside in poverty, but what meager resources they are able to set aside for education are very sensitive to external shocks. Without the proper safety nets to help mitigate or recover from illness or disability, actors can face the double burden of paying out-of-pocket for treatment and foregone pay due to absence. Furthermore, left unprotected by labor regulations, informal employees are at a continually higher risk of being fired than their formal counterparts. Informal households also suffer from an inability to access many of the avenues available to actors in the formal sector (or informal sector in developed countries) to offset the cost of education, such as employee training or sponsorships (Jütting and Laiglesia 2009).

In their model of informality and long-run growth, Docquier et al. (2014) show that, relative to the costs of education, the perceived benefits or incentives for educational investment by informal households may actually be reduced through two mechanisms:

> *On the one hand, since the informal sector absorbs a large share of the unskilled labor force, the supply of unskilled workers to the formal sector is reduced, leading to a smaller skill premium. On the other hand, the occurrence of child labor is facilitated by the existence of the informal economy. Faced with lower skill premia and easier access to child labor, altruistic parents tend to choose less schooling for their children.* (2)

Results from their model are in line with empirical evidence regarding low productivity and increases in child labor in societies with high informality, as well as accounting for the previously unexplained low skill premia found in developing countries. Docquier et al. (2014) conclude that a low level of human capital accumulation sustains a large, unproductive, informal sector and ultimately prevents income convergence towards that of more developed nations.

Oviedo et al. (2009) reaches a similar conclusion, though she arrives at it from the idea that one of the heaviest social burdens of survivalist economies may be the individual's point of view that they, and their children, are powerless to escape a survivalist existence. This perspective, in turn, discourages investment in human capital for themselves or their children, maintaining marginalization indefinitely. The result across multiple generations may actually be an increasing number of informal actors in low-productivity and vulnerable states. While a large portion of the cost of this marginalization may continue to be paid by those in the survivalist economy, negative spillovers could reach the rest of society in the form of reduced overall productivity.

In both the cost-benefit and belief-system stories, intervention may be required to help survivalists escape the poverty trap. Policies promoting investment in human capital, such as by reducing the cost of education, or providing some security in mitigating the potentially disastrous effects of injury or illness, aim to break the positive (and vicious) feedback loop of informality with poverty and low productivity. Model results from Docquier et al. (2014) conclude that education subsidies are the most cost-effective policies for avoiding the poverty trap; in the long-run, however, or in lower-growth scenarios, combining educational intervention with wage subsidies for low-skilled workers to incentivize entrance into formality is preferred.

Drivers of entrepreneurial informality

Today's developed countries have grown out of most of the survivalist pressures that informal actors in less developed countries face. They universally provide at least a basic level of social protection and boast a greater dispersion of human capital. Therefore, developed countries have a lower propensity for informal/formal labor segmentation. Their formal sectors can satisfy the basic needs of most economic actors, and business regulations are not so burdensome as to make actors choose between formality and survival. And yet, informality remains. This indicates the existence of costs associated with formality, which at least some actors are unwilling to pay, even if their survival is not threatened.

Most of the time, the decision to move into the informal sector is made by considering the costs associated with undesired interaction with the state against the benefits it supplies, and while many non-state factors can also weigh on the decision, the vast majority of literature focuses on a set of drivers involving the weight of state intervention. It is recognized, of course, that these same factors also affect developing countries and augment the survivalist perspective in them.

Of the factors normally associated with the cost of remaining formal, and in turn with the overall size of the informal sector, Schneider and Enste (2000) list the following:

> *…the rise of the burden of taxes and social security contributions; increased regulation in the official economy, especially of labor markets; forced reduction of weekly working time; earlier retirement; and the decline of civic virtue and loyalty towards public institutions combined with a declining tax morale.* (82)

While this list is certainly not exhaustive, it does cover the core variables used in most models of the informal economy, in particular the magnitude of tax burden and extent of government regulation, which we will focus on more closely. Although the distinction is not always a very sharp one, it can be useful to further subdivide the drivers of entrepreneurial informality into those that generally push actors away from formality and those that tend to pull them into formality.

Push factors from formality to informality: Taxation, regulation, and corruption

Firms, believing themselves overburdened by taxes and social security payments can choose to evade them by staying or going 'underground', and in doing so are rewarded with a cost advantage and therefore higher earnings relative to those who still remain in the formal sector. Actors are faced with a simple cost-benefit question, one made easier to answer as the burden of taxation increases. Schneider and Enste (2000) claim that in "almost all studies, one of the most important causes of the increase of the shadow economy is the rise of the tax and social security burdens," and they cite over a dozen authors in their support (82). They describe the relationship as:

> *The bigger the difference between the total cost of labor in the official economy and after-tax earnings (from work), the greater the incentive to avoid this difference and to work in the shadow economy. Since this difference depends broadly on the social security system and the overall tax burden, they are key features of the existence and rise of the shadow economy.* (Schneider and Enste 2000: 82)

While this relationship is the most commonly reported, some studies have raised doubt about the claim. Friedman, et al. (2000: 460) reported finding "no evidence that higher direct or indirect tax rates are associated with a larger unofficial economy. In fact, we find some evidence that higher direct tax rates are associated with a smaller underground sector." However, similar to other lines of inquiry, they found that the relationship lost significance after controlling for per capita income as a proxy for better government efficiency and higher tax rates. The resulting narrative they use to explain this observation is that:

> *…when faced with onerous bureaucracy, high levels of corruption, and a weak legal system, businesses hide their activities 'underground'. Consequently, tax revenues fall, and the quality of public administration declines accordingly, further reducing a firm's incentives to remain 'official'… Richer countries have both higher tax rates and a smaller unofficial economy. Across the countries in our sample, the incentive to go underground to dodge higher tax rates is outweighed by the benefits of remaining official when tax rates are higher. This is probably because, at least for this set of countries, higher tax rates generate revenue that provides productivity enhancing public goods and a strong legal environment*. (Friedman, et al. 2000: 460, 476)

The causal pathway described by Friedman et al. (2000) illustrates the difficulty in disentangling and then interpreting results from a model of interdependent variables. A representation of taxes is core to most conceptualizations of the informal economy in developed countries, though its proximity to both important push and pull factors can make interpretation of its actual effect on informality somewhat statistically ambiguous. Nevertheless, after consideration of these points Schneider and Enste (2000: 83) return to their well accepted notion that, *ceteris paribus*, "higher indirect tax rates and higher marginal income tax rates tend to raise the amount of labor and goods bought and sold in the underground sector."

Another very commonly referenced class of determinants is the overburden of government regulation. Many different variables are used, somewhat interchangeably, to explore its relationship with the informal economy. Studies that incorporate regulatory variables often look at the financial or time commitments necessary to start or register a business, the ease in which employers are able to fire employees, or indicators measuring the quality of property rights (de Soto 1989 and 2000), or the overall freedom firms have to conduct business. One important study in this area, Loayza et al. (2005), finds a significant positive relationship between regulation burden on the size of the informal sector (that is, both grow or decline together). Pursuing this inquiry further, they determine that through this established relationship, overly burdensome regulations have a significant negative impact on economic growth, whereby,

> *…if a typical developing country were to decrease its product-market regulation to the median level of industrial countries (that is, from 0.51 to 0.17) while maintaining its level of governance (equal to the median of developing countries, 0.37), then its annual growth rate would rise by about 1.7 percentage points.* (Loayza et al. 2005: 8)

Following a similar logic, corruption, or bribes in particular, can be thought of as an indirect tax that entrepreneurs pay in order to avoid an excessively burdensome regulatory environment. Higher levels of corruption may also indicate higher levels of tolerance officials have toward unreported economic activity (Torrini 2005). Andrews et al. (2011: 30) explain that "This reflects the idea that in countries where there are fewer opportunities of evading taxes, higher taxation will result in a lower rate of self-employment to the extent that the income of self-employed workers is more sensitive to individual effort – and thus more responsive to tax rates – than wage workers."

This nexus of taxation, regulation, and corruption receives a good bit of attention throughout the literature. One interesting line explores a positive (but vicious) feedback loop that links the erosion of the functioning of the regulations, illegitimacy of government institutions, and an increase in informality. In this scenario, Centeno and Portes (2006: 34) describe how:

> *…attempts by states, weak and strong, to impose themselves on civil society by implementing pervasive controls backfire, leading to self-reinforcing circles that negate the intent of the rules. In the case of weak states, such as those of Latin America, state protection and resources—including access to predictable, legal transactions—are appropriated by a minority, while the rest of the population is left to fend for itself through widespread violation of the law.*

At the same time, just the opposite (that is, a virtuous, self-reinforcing dynamic) is possible when, "limited regulation of private activity by a capable state is coupled with widespread legitimacy of existing rules among the citizenry. In these instances, society itself becomes an enforcer and guarantor of the rule of law." (Centeno and Portes 2006: 35)

Pull factors from informality to formality: Government transfers to households and spending on R&D

The other side of the equation actors face regarding their position in the economy pertains to the benefits they receive by deciding to continue or initiate their engagement with the state and its rules. Many factors will influence the assessment of such benefits by each household or firm actor. For a firm these include its capital or labor intensity and needs, whether it seeks growth and expansion, and its network of relationships with other informal or formal actors.

For an individual, one major benefit comes from the social programs and safety nets provided by the state. When these benefits are tied to the formal economy, say through program participation by employers, there are strong incentives to either remain, or move into, the formal sector (Oviedo et al. 2009). Schneider and Enste (2000: 86) however deny that social welfare transfers can offer any possible returns to formality, evoking the classic free-rider claim that "such a system provides disincentives for individuals receiving welfare payments to even search for work in the official economy, since their overall income is higher if they receive these transfers while working in the underground economy.". The few studies that evaluate the impact that various social transfer programs have on informal/formal labor supplies offer support for both arguments.

Nevertheless, evidence suggests that the manner in which these benefits are handled and dispersed can play an important role in whether they pull people into the formal sector or push them away. This is not to suggest, however, that there is an understood best practice for incentivizing formality thorough social benefits. In examining the impact of conditional cash transfer programs in Mexico, Bosch et al. (2012) find no evidence that the programs had a dependence effect, measured as increasing unemployment or informality, but found significant evidence of a labor shift from informality to self-employment. On the other hand, de Brauw et al. (2013) found that Brazil's Bolsa Familia program incentivized a shift of around 8 hours per week per household member from the formal to informal sector. While the impact of welfare programs on formal employment may not be generalizable, one can conjecture that, all else being equal, household transfers, when passing through formal actors, should incentivize participation in the formal sector.

There is, however, a more fundamental issue with respect to pull factors for the formal economy. Societies wanting to move from lower to higher income levels, and from higher to lower informality in the process, build and use public resources to do so. Strengthening human capacity is one of the key targets, as is building physical infrastructure. Increasingly, a focus of countries wanting to avoid middle-income traps is on creating a knowledge society (Agénor and Canuto 2012; Kanchoochat and Intarakumnerd 2014). Knowledge societies benefit from generalized human capacity but also from particular kinds of it, including tertiary education, often with specialized attention to science and engineering. They also benefit from research and development expenditures that again strengthen capacity of individuals and also of firms. They frequently incubate those firms and/or connect them to educational systems. Although the literature on informality gives limited attention to this, in our own analysis of drivers of informality we explored various measures of the knowledge society; but we did not find that they added significant explanatory power, even in combination with some of the other variables discussed above. One partial exception was R&D expenditures. As Tho (2013: 8) says, R&D is "essential for facilitating the transition from a labor-surplus to a labor-shortage economy, the transition from input-driven growth to TFP-based growth, and for upgrading the industrial and export structure to high-skill and technology-intensive products."

It is often difficult to know if the relationships we find and explain below in our own empirical analysis of correlations between both social support programs and R&D, on one side, and lower levels of informality on the other, are capturing a true causal connection among these variables. Because we find them even when we control for years of education, however, we can be quite confident that they do not simply reflect the impact of the level of socio-economic development on both sides of the relationship (therefore being a spurious or false correlation). It is possible, however, that both of the pull factors we have discussed, higher government transfers to households and greater R&D spending, to some degree reflect a broader improved quality of governance overall and focus on the kind of economic and social transformations that do pull people into formality with both opportunities and capabilities. We strongly suspect that both governmental social support for households and R&D spending combine some true pulling power themselves with an element of proxy representation of such strong and effective government.

It is appropriate now to turn to our statistical analysis of drivers, informed by the preceding conceptual and theoretical analysis of the factors that determine the balance between informality and formality across a wide range of countries.

Drivers in IFs

We selected the variables used to represent the drivers of the informal economy in the International Futures model based on the following criteria:

1. The literature identified them as having relative high importance.
2. When combined they cover a large portion of types of drivers "clusters."
3. They represent a variety of tangible policy interventions.
4. The model they specify is significant and robust.

Overall, the final model should contain drivers that both capture the strong long-term relationship between human development and informality as well as the drivers that relate both to the motivations and choices of firms and individuals and to the levers or policies of governments. These drivers can be clustered into the following categories:

1. *Developmental drivers* are the root cause of involuntary informality: low human capital. In developing countries with higher levels of survivalist informality, developmental drivers may dominate. Even in more developed economies, where informality tends to exhibit different characteristics, a deeper developmental driver may retain importance through its impact on the secondary drivers of informality.

2. *Secondary drivers* are the push and pull factors that affect whether an actor will choose to enter, stay in, or opt-out of the formal economy

The variables analyzed for inclusion in the IFs model of informality

In building the formulations to represent informality in IFs, we explored a wide range of possible variables before selecting those actually used. We review some of those below.

Developmental driver

Education Years: *Average years of educational attainment of adults, age 15 and older (Barro and Lee 2010).* Educational attainment enters the model as a deep developmental driver of informal labor share. In most models from the literature, per capita income appears as a control variable, occupying a similar role. In giving preference to education we more directly capture the human capital obstacle that faces informal actors in survivalist situations. This choice will also allow us to construct scenarios with more tangible and relevant policy options. A second, statistical, justification for the choice of education years is that relationship between education years and informal labor share is stronger than that using per capita GDP. Education years has saturating character (when societies achieve high levels of education the rate of further growth slows), meaning that as a driver it should tend to impart asymptotic like behavior to informal labor share; the evidence does suggest that the informal economy is present in even the most developed countries.

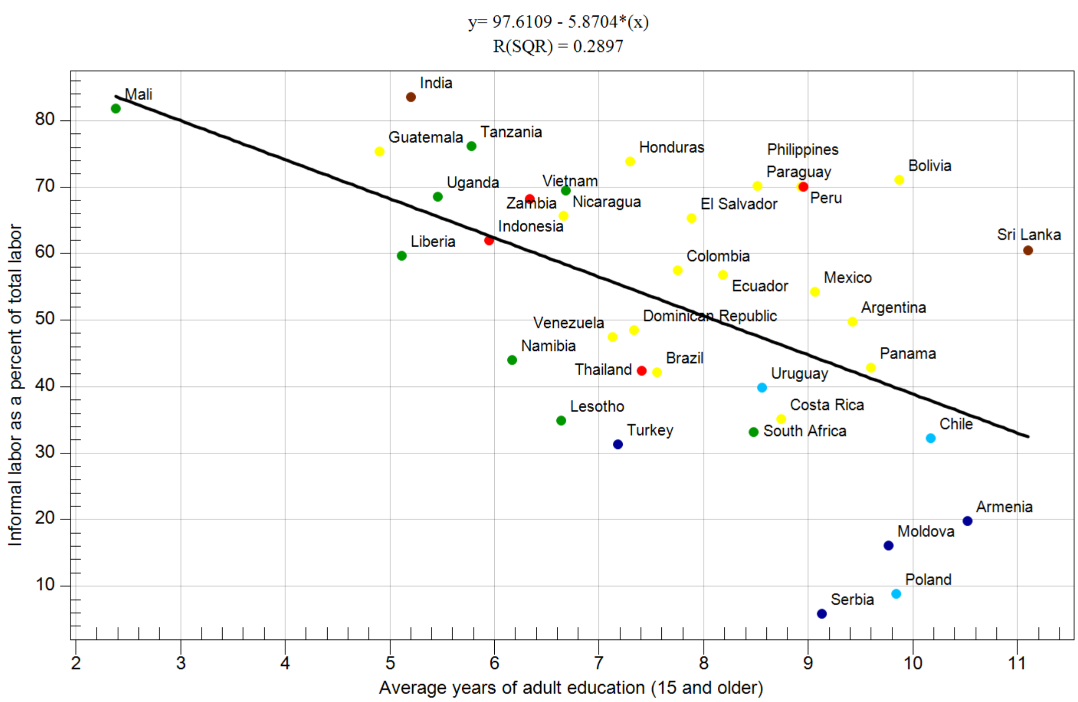

Figure 4.3. Cross-sectional graph between informal labor (as a percent of total labor) and average years of adult education (15 and older)
Source: IFs version 7.14

GDP per capita at PPP (2011 USD): *From the World Bank's World Development Indicators (WDI).* As stated above, in the literature GDP per capita often appears as a control variable. In this project it can be thought of as another developmental driver, however it is rarely found as a persistent explanatory variable due to its high correlation with education years.

Push factors

Government Regulation Index: *A composite index combining indicators measuring bureaucratic obstacles to starting and maintaining a business.* This index combines the following five separate regulatory and overall business environment indicators and indices from their standardized score after correcting for direction (higher represents greater levels of regulation):

- The cost of starting a business from the World Bank's Doing Business database
- The number of procedures needed to start a business from the World Bank's Doing Business database
- An index capturing the ease of firing an employee from the World Bank's Doing Business database
- A measure of business freedom from the Heritage Foundation
- An index reflecting property rights from the Heritage Foundation.

Tax Rate: *The total share of taxed profits and mandatory contributions from all sources as a percent of commercial profit from WDI.*[23] The tax rate, as it is included in this model, has a positive relationship with the size of the informal labor share (higher taxes are associated with a larger informal share). However, as found in other studies (and mentioned above) significance (and sometimes even direction) of the tax variable can be very dependent on other variables included in the model, particularly those controlling for government regulation and rule of law. In fact, when we tested substitution of a measure of corporate taxes expressed as a percent of GDP, we found that the sign on the tax variable's coefficient becomes negative and significant.

Government Corruption: *From Transparency International's Corruption Perception Index. Measures the perceived level of government corruption.* The name is somewhat misleading as higher values indicate greater government transparency, where a

---

[23] A description of the total tax rate come from the World Bank's Doing Business website:

> *The total tax rate measures the amount of taxes and mandatory contributions borne by the business in the second year of operation, expressed as a share of commercial profit… The total tax rate is designed to provide a comprehensive measure of the cost of all the taxes a business bears. It differs from the statutory tax rate, which merely provides the factor to be applied to the tax base. In computing the total tax rate, the actual tax payable is divided by commercial profit.*

The tax rate appears in the model as its logged transformation.

score of 100 is the least corrupt (the longer-term time series of the measure in IFs is scaled from 0-10). Corruption can be thought of as an indirect tax that entrepreneurs pay in order to avoid an excessively burdensome regulatory environment. Higher levels of corruption may also indicate higher levels of tolerance officials have toward unreported economic activity (Torrini 2005).

Unemployment: *From WDI.* In spite of what intuition might suggest, many studies find an ambiguous relationship between unemployment and informality (Lisi and Pugno 2011; Schneider and Williams 2013; Buehn and Schneider 2012; Enste 2009; Macias 2008). Tanzi (1999) offers one explanation in which for many countries a large portion of the informal economy may be comprised of minors, retirees, housewives, and immigrants, all of whom can be excluded from the official workforce. Due to this conceptual and empirical ambiguity (and because the variable is not forecast in IFs), we do not include unemployment in our statistical analyses.

Pull factors

Spending on R&D: *From WDI: Expenditures for research and development are current and capital expenditures (both public and private) on creative work undertaken systematically to increase knowledge, including knowledge of humanity, culture, and society, and the use of knowledge for new applications. R&D covers basic research, applied research, and experimental development.*[24] We use this indicator as a measure of the knowledge society discussed above, which strengthens the capacity of firms and individuals making them more competitive in the formal sector.

Government to Household Transfers: *The sum total of government transfers to households in the form of pensions and welfare, expressed as a percent of GDP.* Values used to specify this variable were estimated in the International Futures model as the difference between direct government consumption and total government expenditures.

Government Effectiveness: *From the World Bank's Worldwide Governance Indicators,* "Government Effectiveness captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies." Centeno and Portes (2006) discuss the importance of public perception of in establishing a culture of tax-compliance that can weather higher levels of taxation because the payers understand that their contributions are being properly used for the provision of social goods.

[24] Definition taken from World Bank World Development Indicators Database research and development series at http://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS [accessed on 3/25/15].

Structural Variables

GDP Growth Rate: *The annual growth rate of GDP from the WDI.* Structuralists claim a pro-cyclical relationship between the informal and formal economies, such that when one experiences economic expansion or contraction, so does the other (Tahir 2014; Heintz and Pollen 2003). On the other hand, Dualists view the informal sector as a residual sector that should eventually disappear once the formal sector of a country has reached full maturity (Bacchetta et al. 2009). The ambiguous relationship found in both the literature and the data has led us to exclude the variable from subsequent statistical analysis.

Gini Index: *A measure of income distribution, where a score of 1 represents complete inequality and 0 represents complete equality*. *From WDI.* Some studies have found a strong positive relationship between income inequality and the size of the informal economy, hypothesizing that higher levels of inequality can lead to social alienation and undermine institutional quality, making the informal sector more appealing for poor individuals who lack the productivity to be competitive in the formal sector (Ahmed et al. 2005). Furthermore, Chong and Gradstein (2004) make the point that inequality will in turn increase as people lose or give up benefits as they move into the informal sector. We tend to see this measure mostly as a result of changes in the level of informality, not a cause of it.

Population Growth Rate: *The annual growth rates of the population from WDI.* Dualists theorize that if population growth exceeds the number of available formal sector jobs, the informal sector will grow to absorb the excess labor (Chen 2012).

Statistical summary

Variable	Name in International Futures	Indicator	Designation	Exp. Sign	First Year and Last Year	Countries	Source	Studies
Educational Attainment	EdYearsAge15Total	Average years of education completed for adults 15 and older	Developmental	-	1950, 2010	144	Barro and Lee	Loayza (2007)
GDP per Capita	GDP2011PCPPP	GDP per capita in $2011 at purchasing power parity	Developmental	-	1960, 2013	186	World Bank WDI	Loayza, et al. (2005)
Business Regulation	GovWBDoingBusinessCostofStarting GovWBDoingBusinessProcedures GovWBDoingBusRedundancyIndex GovHeritageBusinessFreedom GovHeritagePropertyRights	Composite index measuring bureaucratic obstacles to starting and maintaining a business. Scale 0—10, higher = more regulation	Push	+	2004, 2010	177	WB Doing Business and Heritage Foundation	Loayza, et al. (2005)
Tax Rate	GovWBDoingBusTotalTaxRate	Total share of taxes from all sources as a percent of commercial profit	Push	+	2005, 2015	186	WB Doing Business	Schneider and Enste (2000)
Corruption	Corruption	Index measuring perceived level of government corruption, scale 0—10, higher = more transparent/less corruption	Push	+	1995, 2011	181	Transparency International	Torrini (2005)
Unemployment	SOFILaborUnemployment	Total unemployed workers as a percentage of the total labor force	Push	+ / -	1985, 2006	145	World Bank WDI	Buehn and Schneider (2011)
Spending on R&D	RANDDEXP	Total spending on research by public and private sector as a percentage of GDP $2011 at MER	Pull	-	2010, 2010	132	Computed in IFs (2010 base value)	None
Government Transfers to Households	GOVHHTRN	Total government pension and welfare transfers to households as a percent of GDP $2011 at MER	Pull	-	2010, 2010	186	Computed in IFs (2010 base value)	Bachetta et al. (2009)
Government Effectiveness	GovernanceEffect	Index measuring perceived effectiveness of government institutions. Scaled -2.5—2.5,	Pull	-	1996, 2012	185	World Bank WGI	Torgler and Schneider (2007)

		higher = more effective						
GDP Growth Rate	GDPR (growth rate calculated externally)	Annual growth rate of GDP $2011 at MER	Structural	+ / -	1960, 2016	186	IMF WEO	Lisi and Pugno (2011)
Gini Index	Gini	Index measuring income inequality. Scaled 0—1, 0 = total inequality, 1 = total equality	Structural	+	1978, 2012	156	World Bank WDI	Chong and Gradstein (2004)
Population Growth Rate	POPR (growth rate calculated externally)	Annual growth rate of population	Structural	+	1961, 2006	182	Calculated from UNDP	Ela (2013)

Table 4.1. List of variables, including their causal role and expected sign in the mode, the data source, and a predominant study of the informal economy in which they are found

Note: the sign on Corruption indicates the expected influence of corruption itself, not the sign of the variable's coefficient, which is expected to be opposite (recall that higher values the Corruption variable indicate greater transparency). WB stands for World Bank, WDI stands for World Development Indicators, and WGI stands for Worldwide Governance Indicators.

Table 4.1 lists the drivers that were selected for more extensive analysis and possible inclusion in the model of the informal economy in IFs after exploration of tens of candidates from the literature and our own analysis. The table also indicates their conceptual role in the model (developmental driver, push factor, and pull factor), their expected sign, the data source, and a predominate study in which they appear. Summary statistics for the potential universe of driving variables are listed in Table 4.2. Table 4.3 shows a correlation matrix of the principal measures of informality used in this study, alongside the strongest candidate drivers of informality discussed above.[25]

Variable	Obs	Mean	Std. Dev.	Min	Max
Education	144	8.16	2.63	1.81	13.09
GDP per Cap	186	16,711.20	18,655.31	584.37	127,541.30
Business Reg	169	4.90	0.64	3.48	6.64
Tax Rate	182	40.62	21.44	7.40	216.50
Corruption	176	4.26	1.99	0.80	9.10
Unemploy	165	10.14	8.63	0.50	59.50
Spend R&D	131	0.81	0.94	0.01	4.40
Gov Trans to HH	186	10.72	8.53	0.00	33.11
Gov Effect	181	2.45	0.97	0.84	4.71
GDP GR	186	0.04	0.03	-0.11	0.14
Gini Index	158	40.21	9.22	24.70	65.77
Pop GR	186	1.54	1.55	-1.10	11.89

Table 4.2. A statistical summary of the most recent values for the model's potential driving variables

Note: Refer to Table 4.1 above for most recent year used for each variable.
Source: IFs version 7.14

	Informal Labor	Informal GDP	Inf Lab Out Inf	Education	GDP per Cap	Business Reg	Tax Rate	Corruption	Unemploy	Spending R&D	Gov Trans to HH	Gov Effect	GDP GR	Gini Index	Pop GR
Informal Labor	1.00														
Informal GDP	-0.16	1.00													
Inf Lab Out Inf	-0.62	0.53	1.00												
Education	-0.54	-0.30	0.38	1.00											
GDP per Cap	-0.43	-0.77	0.19	0.78	1.00										
Business Reg	0.46	0.47	-0.25	-0.62	-0.61	1.00									
Tax Rate	0.37	-0.20	-0.41	-0.07	-0.20	0.33	1.00								
Corruption	-0.43	-0.76	0.01	0.66	0.69	-0.77	-0.19	1.00							
Unemploy	-0.53	0.17	0.56	0.12	-0.03	0.07	-0.13	-0.06	1.00						
Spending R&D	-0.47	-0.74	0.26	0.54	0.53	-0.54	0.08	0.69	-0.07	1.00					
Gov Trans to HH	-0.60	-0.43	0.27	0.67	0.68	-0.41	0.07	0.52	0.01	0.48	1.00				
Gov Effect	-0.43	-0.76	0.08	0.73	0.78	-0.80	-0.19	0.93	-0.10	0.71	0.58	1.00			
GDP GR	0.32	0.45	-0.24	-0.43	-0.42	0.29	-0.07	-0.26	-0.09	-0.33	-0.43	-0.33	1.00		
Gini Index	0.21	0.16	-0.07	-0.28	-0.25	0.30	0.07	-0.20	0.12	-0.42	-0.39	-0.26	0.14	1.00	
Pop GR	0.70	-0.14	-0.45	-0.44	-0.22	0.26	-0.16	-0.20	-0.18	-0.23	-0.47	-0.27	0.32	0.20	1.00

[25] Explanatory variables at this point have been transformed for normality: Tax Rate and GDP per capita have undergone log transformations, and Government Spending on R&D has undergone a square-root transformation.

Table 4.3. A Pearson product moment correlation matrix of the measures and driving variables of informality
Note: Refer to Table 4.1 above for most recent year used for each variable.
Source: IFs version 7.14

Building the models

In this section we develop three statistical models that will be used in our forecasts of the informal economy: (1) a model of the size of the informal labor force, (2) a model of the distribution of informal labor residing either inside, or outside the informal sector, and (3) a model of the informal economy expressed as a percent of official GDP. As mentioned above, modeling efforts will focus especially heavily on informal labor because of its relative strong empirical base and its role as a prerequisite for the existence of an informal economy. Therefore, we first develop a model of the size of the informal labor force, and then consider the output as a potential driving variable for the subsequent two models.

Informal labor as a percent of total labor

Starting with the full set of variables we began to explore best sub-set specification options. Initial results indicated consistent significance and stability in the education, regulation, taxation, and government transfers variables. By comparing the relative quality of the models using Akaike information criterion (AIC) and the Bayesian information criterion (BIC) two models were identified as best subsets from the universe of explanatory variables. Table 4.4 shows the regression outputs of these two models (Model 1 and Model 2).

While Model 1 reports a higher R^2, the variables measuring spending on R&D and inequality (Gini) are found to be insignificant. After conducting the first round of a backwards-stepwise regression algorithm we remove spending on R&D and are left with the specification found in Model 2, in which the variable measuring inequality is also is found to be insignificant. The resulting Model 3 is comprised of four variables, education, business regulation, tax rate, and government transfers to households, each with significant coefficients and with the expected sign.[26]

[26] A more technical description of the model specification process can be found in Appendix A4.

	Model 1	Model 2	Model 3
Education	-3.514**	-2.676*	-2.510*
	(-3.32)	(-2.53)	(-2.37)
Business Reg	10.86**	12.51**	12.60**
	(3.02)	(3.45)	(3.45)
Tax Rate	10.89*	11.08*	10.60*
	(2.55)	(2.57)	(2.45)
Gov Trans to HH	-1.402***	-1.737***	-1.810***
	(-4.17)	(-6.04)	(-6.38)
Gini	0.347	0.236	
	(1.80)	(1.23)	
Spending R&D	-12.58		
	(-1.69)		
Constant	-11.77	-24.14	-12.61
	(-0.53)	(-1.07)	(-0.61)
Observations	34	37	37
R^2	0.827	0.779	0.768

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.4. Ordinary least squares regression estimates for the top three models

Note: A more technical description of the model specification process can be found in Appendix A4. Refer to Table 4.1 above for most recent year used for each variable.

Source: IFs version 7.14

Table 4.5 shows ordinary least squares regression estimates of each driver against informal labor share using the most recent data point available for each country-variable. Allowing each variable to enter alone we can see that each is significant and in the expected direction. Overall, the model explains around three quarters of the informal labor size across our country set.

	Model 1	Model 2	Model 3	Model 4	Full Model
Education	-5.870***				-2.510*
Business Reg		18.17**			12.60**
Tax Rate			13.95*		10.60*
Gov Trans to HH				-1.803***	-1.810***
Observations	37	40	41	41	37
R2	0.29	0.208	0.137	0.36	0.768

* p<0.05, ** p<0.01, *** p<0.001

Table 4.5. Ordinary least squares regression estimates for each driving variable (Models 1 – 4) and the Full Model
Note: Refer to Table 4.1 above for most recent year used for each variable.
Source: IFs version 7.14

The coefficient on education in the Full Model indicates that for an increase of one year of average education for people 15 years of age and older, we would expect to see a decrease in informal labor share of 2.5 points. For an increase of 1 percent of GDP (or in the case of Peru, approximately $2 billion in 2014) in government transfers to households by way of pension and welfare, the model suggest a decrease in informal labor of 1.8 points. Focusing on the variables which drive employment out of the formal sector, the model suggests that a one percent increase in the total tax rate would yield a growth in informal labor share of around 0.1 percentage points, and a increase of overall business regulations by 1 standard deviation would yield a 0.33 standard deviation increase in the informal labor share (or stated otherwise, a 1 point increase in the business regulation index could induce an increase of around 12.6 percentage points).

Figure 4.4 shows a cross-sectional plot of the model-predicted country values against the most recent informal labor data. From the graph we can see that the model has underestimated Peru's informal labor share by around 21 points according to the data from ILO-WEIGO, and 17 points according to data provided from INEI (see Table 4.6). For comparison, Chile, Colombia, and Mexico have also been highlighted.

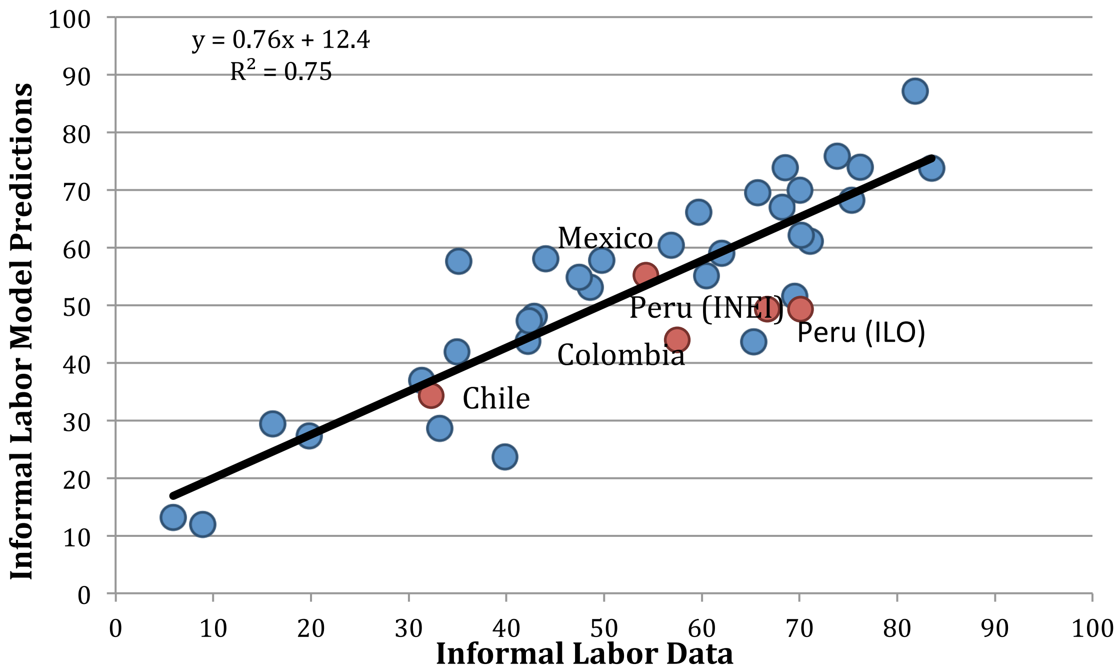

Figure 4.4. Cross-sectional plot comparing the statistical model-predicted values and data for the informal labor share
Source: IFs version 7.14

Country	Informal Labor Data	Informal Labor Model	Educational Attainment (years)	Business Regulation Index	Tax Rate (% Profits)	Govt to HH Transfers (% GDP)
Peru	70.10 (66.7)	49.27	8.93	5.05	36.00	9.56
Chile	32.30	34.30	10.17	4.44	27.90	10.41
Colombia	57.50	43.98	7.75	4.45	75.40	14.26
Mexico	54.30	55.26	9.06	4.90	51.80	7.18

* Values in parenthesis are taken from INEI (2014).

Table 4.6. Comparison of informal labor and its drivers for Peru, Chile, Columbia and Mexico.
Note: the model is initialized with ILO-WIEGO data, but the data from the INEI for Peru has been included in the model's data project file and the model can be initialized to use this alternative measure. The data for the INEI is for the informal sector only, not the entire informal economy. Refer to Table 4.1 above for most recent year used for each variable.
Source: IFs version 7.14

As seen in Figure 4.4 and Table 4.6, our statistical model underestimates the overall share of informal labor in Peru. This underestimation is primarily due to there being other determinants of informality not captured in our model—among the most significant of these may well be the historical path dependency rooted in factors such as culture and past policies. To account for this historical path dependency we initialize our forecasts with values for 2010 based on historical data – a process that allows for the seamless integration of existing data with country-level forecasts driven by an understanding of the global patterns of informal labor.

Informal labor residing inside the informal sector

Using a similar process as applied to the specification of the informal labor share model, we began with the universe of variables identified above, and identified a best subset based on relative quality as indicated by AIC and BIC measures. The resulting best sub-set model (Model 1 in Table 4.7) contains variables that are not significant at the 95% confidence level. Therefore we preformed a backward stepwise selection, removing the least significant variables from the model until all remaining variables were found to be significant, determining that informal labor share alone specified the most comparatively robust and significant model (Model 3).[27]

	Model 1	Model 2	Model 3
Informal Labor	0.464***	0.464***	0.417***
	(5.16)	(5.43)	(4.47)
Corruption	3.021	2.777	
	(1.70)	(1.80)	
Rule of Law	-10.94		
	(-0.41)		
Constant	37.23*	32.66***	44.65***
	(2.64)	(3.76)	(8.33)
Observations	31	32	33
R^2	0.511	0.504	0.392

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.7. Ordinary least squares regression estimates for the best subset model, and its backward stepwise selection derivatives

Note: Refer to Table 4.1 above for most recent year used for each variable

Source: IFs version 7.14

Our final model (Model 3) suggests that for a one percentage point increase in informal labor, the share of informal labor that can be found operating within informal enterprises (or that can be found outside formal enterprises and households) increases by 0.4 percentage points. Since the model for informal labor (as well as the available historical data) suggests a decline in informal labor with

[27] [27] A more technical description of the model specification process can be found in Appendix A4.

increasing levels of development we would expect over time a general trend of informal labor flowing out of the informal sector.[28]

Informal GDP share: NOE and Shadow Economy

Since informal economic production is predicated on informal factors of productivity, namely labor, we consider the informal employment variable a core explanatory variable in our models of informal GDP share (NOE formulation) and the Shadow Economy. With this variable fixed in the model, the best subset of significant variables (after removing the explanatory variables for informal labor and introducing a variable measuring the informal labor residing inside the informal sector as a percent of total employment) also included government corruption and government spending on R&D.[29] Both have the expected sign (the corruption variable appears with a negative coefficient because higher values indicate higher levels of transparency), and are found to be significant at the 95% confidence level.

	Model 1	Model 2	Model 3	Model 4
Informal Labor	0.159	0.133	0.193***	0.187***
	(1.70)	(1.45)	(3.93)	(4.15)
Spending R&D	-2.545	-2.790*	-2.860*	-2.613*
	(-1.97)	(-2.19)	(-2.25)	(-2.15)
Corruption	-1.600	-1.711	-1.593	-1.630*
	(-1.50)	(-1.61)	(-1.52)	(-2.36)
Rule of Law	1.301	3.403	2.620	
	(0.09)	(0.23)	(0.17)	
Informal Labor Split	0.126	0.140		
	(0.71)	(0.78)		
Labor Growth	-63.34			
	(-1.23)			
Constant	12.88	11.55	17.73*	19.36***
	(1.21)	(1.09)	(2.51)	(4.44)
Observations	113	113	113	115
R^2	0.550	0.543	0.541	0.546

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[28] By forecasting informal labor share residing inside the informal sector we are implicitly forecasting the share of informal labor residing outside the informal sector, which can be thought of as the complement of the former measure.

[29] A more technical description of the model specification process can be found in Appendix A4.

Table 4.8. Ordinary least squares regression estimates for all informal economy (NOE formulation) models included in the backward-stepwise regression analysis.
Note:. Refer to Table 4.1 above for most recent year used for each variable.
Source: IFs version 7.14

The model we chose to use for forecasting the informal economy (non-observed economy formulation) share of GDP (Model 4) suggests that for a 1 percentage point increase in the informal labor share, we would expect the informal GDP share to increase by around 0.2 percentage points. With a 1 percentage point increase in spending on R&D (measured as percent of GDP) the model suggests a 2.6 percentage point decrease in informal GDP share, and with a 1 point decrease in corruption (or increase in transparency to stay consistent with the measure) we would expect a 1.6 percentage point decrease in informal GDP share.

	Model 1	Model 2	Model 3	Model 4
Informal Labor	0.0759	0.128*	0.129**	0.102*
	(0.80)	(2.44)	(2.63)	(2.22)
Spending R&D	-2.727*	-2.780*	-2.699*	-2.957*
	(-2.12)	(-2.17)	(-2.22)	(-2.44)
Corruption	-1.651	-1.547	-1.898**	-1.885**
	(-1.55)	(-1.48)	(-2.74)	(-2.71)
Labor Growth	-76.28	-78.31	-75.86	
	(-1.49)	(-1.54)	(-1.51)	
Rule of Law	-3.307	-3.969		
	(-0.22)	(-0.26)		
Informal Labor Split	0.119			
	(0.66)			
Constant	31.98**	37.17***	36.31***	36.28***
	(3.01)	(5.18)	(8.23)	(8.18)
Observations	112	112	114	114
R^2	0.489	0.486	0.490	0.479

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.9. Ordinary least squares regression estimates for all Shadow Economy models included in the backward-stepwise regression analysis.
Note:. Refer to Table 4.1 above for most recent year used for each variable.
Source: IFs version 7.14

Very similarly, the model we chose to use for forecasting the Shadow Economy (Model 4) suggests that for a 1 percentage point increase in the informal labor share, we would expect the Shadow Economy share of GDP to increase by around 0.1 percentage points. With a 1 percentage point increase in spending on R&D (measured as percent of GDP) the model suggests around a 3 percentage point decrease in informal GDP share, and with a 1 point decrease in corruption (or increase in transparency to stay consistent with the measure) we would expect a 1.9 percentage point decrease in informal GDP share.

Historical analysis of informality models

Across the three models developed in this section are seven explanatory variables representing the deep developmental variables that drive human capital, the factors that push entrepreneurs and laborers away from the formal economy, the factors which incentivize activity within the formal sector, and the structural characteristics of the economy that may influence the ways in which wealth and benefits are distributed within a country. Figure 4.5 illustrates these drivers and the direction of their impact on the informality variables.

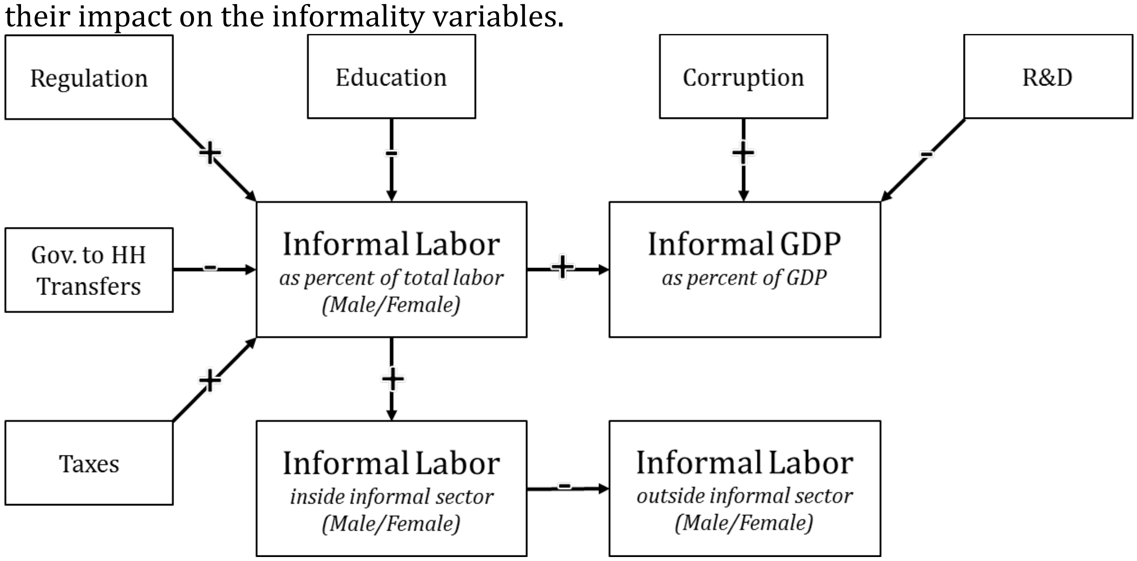

Figure 4.5. Drivers of informality in IFs
Source: authors' conception

Our historical analysis focuses on two primary outputs: (1) the initial historical estimation of levels of informality, and (2) the overall estimated trend relative to the historical trend. Since in forecasting these variables we used the models specified above to evolve measures initialized by actual data, we are primarily interested in the second of these two outputs.

Because the historical data coverage varies greatly between each of these explanatory variables, we had to do a good deal of imputation, and hole-filling, in order to explore what the historical estimates of our forecasted informality measures would be. For some of these variables, such as the average years of education of people 15 years of age and older (Figure 4.6), the longer trend evolution lends itself to a linear interpolation of the 5-year data points reported by Barro and Lee (2010).

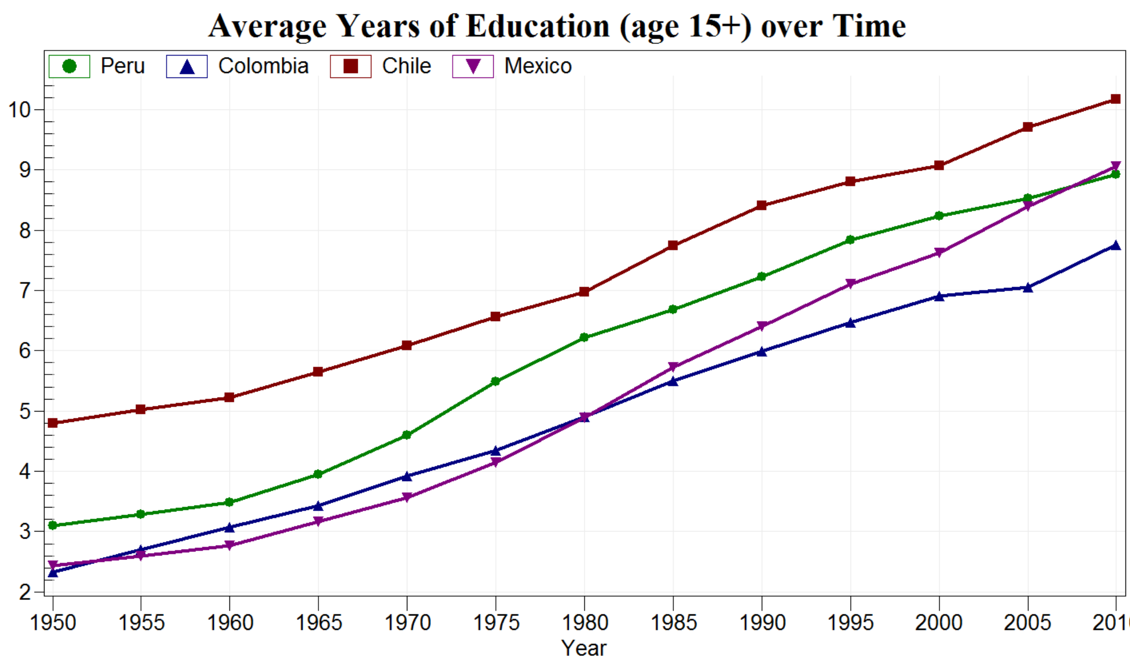

Figure 4.6. The evolution of educational attainment for four countries, 1950—2010
Source: IFs version 7.12

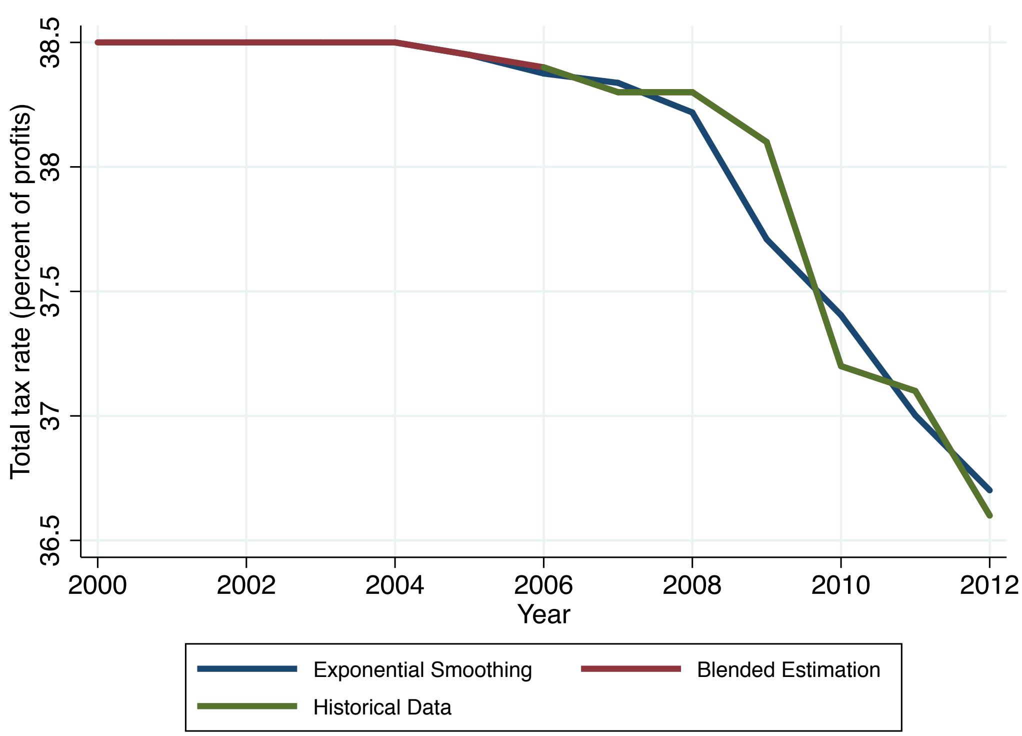

Figure 4.7. Estimating the total tax rate for Peru, 2000—2012
Source: World Bank (2014) and authors estimations

The corruption variable from Transparency International only has two years of data, however we were able to blend it with an older version of the measure to obtain values back to 1998. For other measures, such as the tax rate, which exhibits more of a sigmoidal trend, we used an exponential smoothing technique to estimate additional years for Peru back to 2004 and then held the values constant once they appeared to level out (see Figure 4.7). For others, such as spending on R&D, government transfers to households, and the business regulation index, we held constant over time. The resulting dataset allows us to backcast our forecast variables from the model's current initialization year (2010) back to the year 2000.

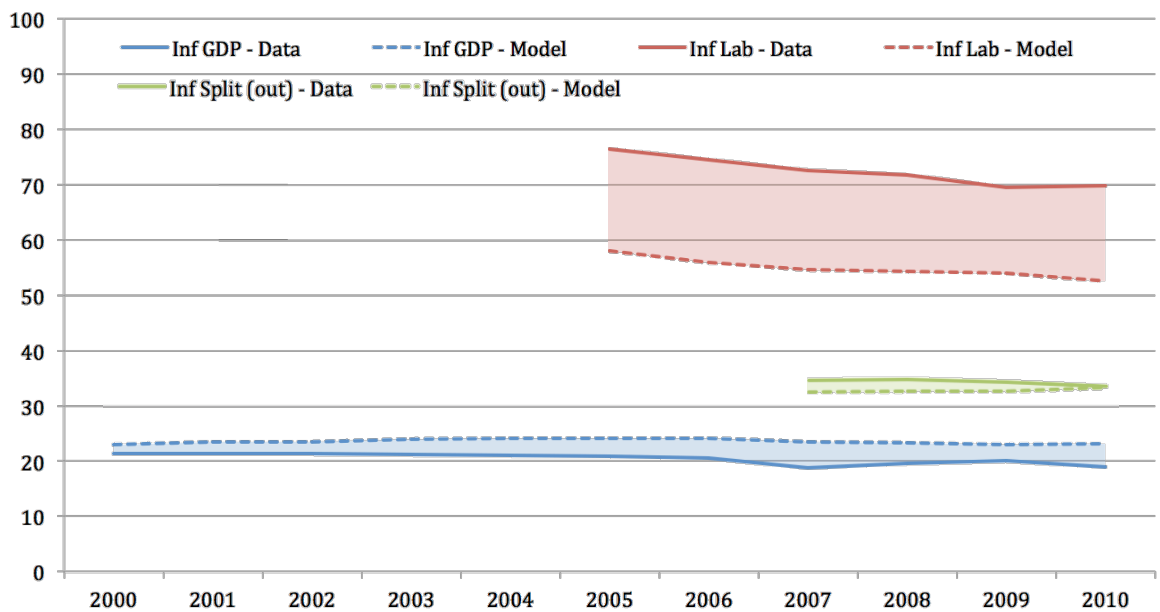

Figure 4.8. Historical data and estimates of informality measures, 2000—2010
Note: Inf GDP is the informal GDP share (in percent of GDP), Inf Lab is the informal labor share (in percent of total labor), and Inf Split (out) is the share of informal labor residing outside of the informal sector.
Source: IFs version 7.14

Originally seen in Figure 4.4, the model underestimates the overall share of informal labor in Peru. Figure 4.8 shows that this underestimation stays relatively constant with an average difference of 17.6 percentage points from 2005 to 2010. Despite this underestimation, the models for informal GDP share and informal labor residing outside the informal sector are very close to actual data. Because our own forecasting initializes values for 2010 based on data, it is the historical ability of our formulations to represent dynamics with time appropriately that is critical to us. Figure 4.9 chains the model backcast with the 2010 initialization value of each variable. Particularly for the informal labor share and the informal GDP share, this graph indicates that the model backcast has captured the appropriate trend for informal labor (and for the other models as well).[30]

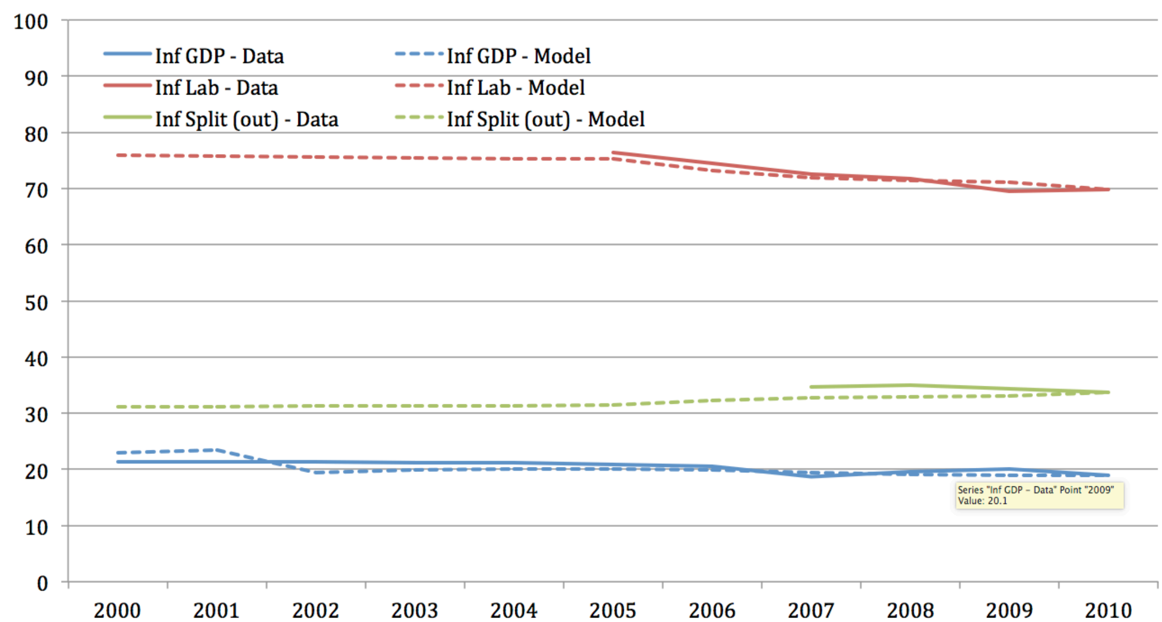

Figure 4.9. Historical data and estimates of informality measures, 2000—2010
Note: The model estimates for each variable have been chained to their respective 2010 historical value.
Source: IFs version 7.14

Now that we have looked at the determinants of informality, particularly as they exist in the IFs model, we turn to the forward linkages, or consequences, that the informal sector may give rise to for a country.

[30] Panel regressions using the data and model for informal labor and GDP confirm this strong and significant relationship over time. It is also important to note that the relatively static trends seen in the informal labor and informality split variables can be attributed to the hole-filling procedures mentioned above.

5. Impacts of Economic Informality

In an early and very influential study of informal economies, Hernando de Soto and his colleagues explored in considerable detail the costs and benefits of both formal and informal activities for the participants in them. Turning then to the larger societal impact of informality, de Soto concluded that the division of society into the two components has:

> *...adverse effects on the economy in general, the main ones being declining productivity, reduced investments, an inefficient tax system, increased utility rates, limited technological progress, and a number of difficulties in formulating macroeconomic policy. (de Soto 1989: 173)*

Subsequent studies have not significantly changed this listing, and it will guide much of our discussion in this section.

Much other literature has identified and explored, almost always qualitatively and often even impressionistically, the economic impacts of informality. Nearly all such studies have argued that informality has a net negative influence on the economy, but the empirical evidence is often insufficient to establish precise causal connections, much less to identify their character and strength. Overall, there appears to be a better understanding about the causes of the informal economy than the costs associated with it. Productivity and tax compliance, two of the elements highlighted by de Soto, have nonetheless often been special foci in the literature. In this section, we will focus on them and discuss the basis for our modeling of forward linkages to them.

Looking beyond effects on the economy and government revenues, there are other, broader social costs of informality including, as de Soto also discussed, increased corruption, less well-developed infrastructure, and reduced progress in human capital development. Many of these broader impacts will automatically appear in our forecasting with the International Futures system as a result of its existing representation of forward linkages from the economy and government revenues to these same variables.[31]

---

[31] There could, of course, be important effects of informality on such variables independent of the paths via economic performance or government revenues—for instance, corruption levels could logically be higher in a society with much informality than in another country, at the same level of GDP per capita, but having low levels of informality. Our modeling even of these two key forward linkages, however, and of their secondary consequences, greatly advances quantitative analysis of the impacts of informality.

Informality's linkage to productivity and economic growth

Theoretical arguments at the micro level claim a number of potential sources of "drag" of informality on economic performance, such as the informal sector's inability to achieve scale and its diminished potential for human capital accumulation.[32]

The smaller a firm is, the easier it is for it to remain hidden from audits and other tax enforcement mechanisms. This may be why the micro-firm is the norm, if not the rule, in the informal economy. In order to hold on to the comparative cost advantage of not paying taxes, these firms must remain undetected, and thus remain small. That micro-size can make it impossible to achieve productive scale.

Baily et al. (2005: 18) claim that this limbo is "a much larger barrier to growth than most policymakers in emerging—and developed—economies acknowledge." In benchmarking the informal sector's drag to a country's productivity by way of a counterfactual, Baser et al. (2006) explore the productivity gains that could result from this uneven playing field. The report begins by claiming that 50 percent of the productivity gap between the US and Turkey can be attributed to the informal economy. In the evaluation of a scenario in which all structures that give small firms a cost advantage by hiding in the informal sector are removed, they find that Turkish productivity could rise from levels currently 40 percent that of the US to an impressive 70 percent. Perry et al. (2007) abates this optimism with the finding that limitations to firm growth are more a function of their size rather than conditions of formality.

The persistently low levels of human capital accumulation of the informal sector may also hamper aggregate growth. As mentioned above, the model in Docquier et al. (2014) corroborates this claim in the case of poverty-based informal sectors. They find evidence for subdued growth in these countries where incentives to invest in human capital, specifically education, is outweighed by the short-term prospects of higher marginal salaries and the value of child labor. They conclude that, "although informality serves to protect low-skilled workers from extreme poverty in the short-run, it prevents income convergence between developed and developing nations" (34).

Despite these compelling theories, strong empirical evidence of the connection is scarce. In one empirical study exploring the causes and consequences of the informal sector in Peru, Loayza (2007: 7-8) finds that "The harmful effect of informality on growth is not only robust and significant, but its magnitude makes it also economically meaningful – An increase of one standard deviation in any of the informality indicators leads to a decline of 1-2 percentage points in the rate of per capita GDP growth." He arrives at this conclusion through the simple regression of

[32] See Oviedo et al. (2009) for a survey of these theories.

various measures of informality against the GDP growth rates for countries over a 20-year span.

Such existing statistical evidence aside, the claim that informality hinders growth still attracts a good bit of criticism, most of which demonstrates that with the inclusion of other standard growth determinants, such as education, the relationship loses its significance. Perry et al. (2007: 173, 175) observes, however, that that this:

> *...could be due to the fact that many of the standard drivers of growth are also likely to affect informality—for example, low levels of human capital or institutional quality leading to both lower growth and higher informality—and it is difficult to separate their direct growth effects from those that operate through larger informal sectors*...

Nevertheless, he concludes that the informal economy has a negative effect on aggregate growth. Thus a common denominator of most studies informal sectors across the world is a belief in their low productivity.

Schneider and Williams (2013) add another dimension to the analysis, however, suggesting that for the most developed countries, informal labor productivity approaches that of the formal sector. This leads us to speculate that informal labor productivity might increase as a function of development.

To help us assess the productivity loss due to informality, we compared the share of GDP that is informal with the share of the labor force that is informal. Should those shares be equal (if, for instance, 50 percent of a country's GDP were informal and it were produced by 50 percent of the labor force, also informal), we could conclude that informality created no obvious productivity loss.

Figure 5.1 shows the ratio of informal GDP share to informal labor share as a function of educational attainment level of countries.[33] The linear regression indicates a statistically significant relationship in the expected positive direction meaning that more developed countries do, indeed, have roughly comparable productivity in informal and formal sectors. Most countries, however, and all with educational attainment of 7 years or less, fall below the 1-to-1 informal GDP/formal labor productivity parity level.[34] Peru's value is 0.75, suggesting a loss of about 25 percent in the productivity of each worker in the informal economy relative to those in the formal sector. Section 5 will explain how we have taken the ratios of Figure 5.1 into our forecasting formulations for linking informality to economic growth.

---

[33] We have data for only 41 countries.

[34] Three formerly communist countries show potentially greater productivity in the informal than in the formal economy. Although this might reflect measurement problems, it could also suggest the continuation of a pattern that was often said to characterize countries under communism, namely a very inefficient formal sector and more efficiency in small private and often informal enterprise.

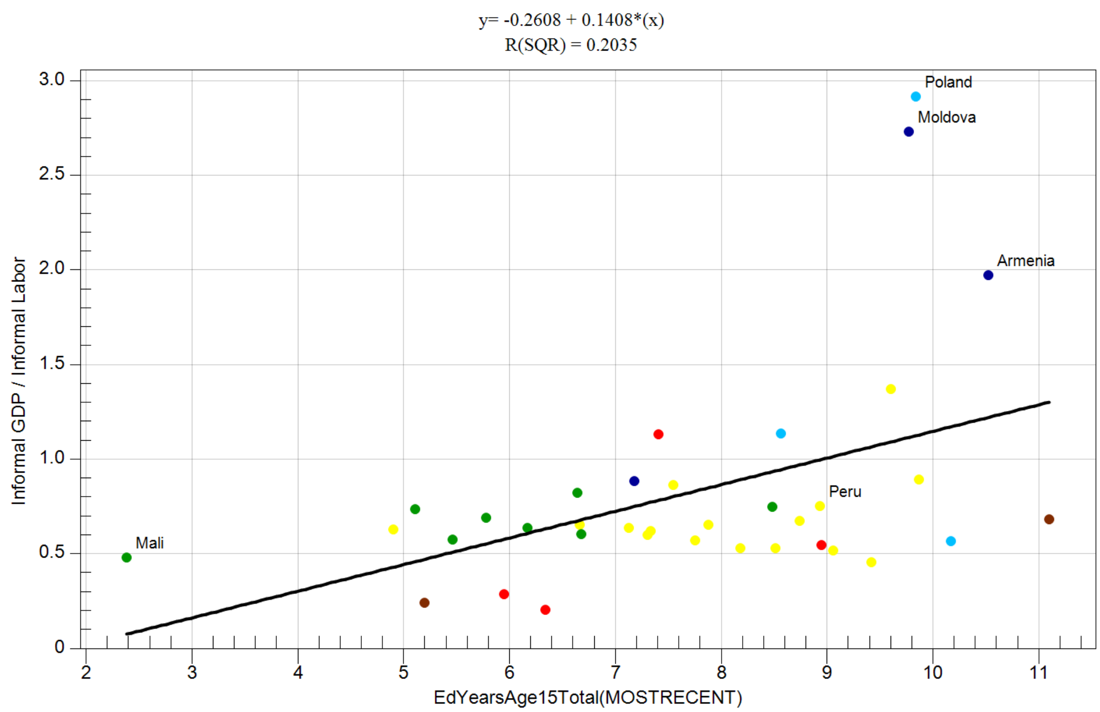

Figure 5.1. The ratio of informal GDP share to informal labor share as a function of educational attainment of adults 15 years of age and older
Source: IFs version 7.14, originally Elgin and Oztunali (2012); ILO (2012). Education data from Barro and Lee (2010)

Informality's linkage to taxes, social provision, and inequality

Moving beyond the impact of informality on productivity and GDP growth and to consideration of its impact on government tax collection and revenues, the basic relationship is quite a simple one—the informal sector seeks to avoid taxes, and that reduces the revenues that the government raises, often causing government to raise tax rates, which both burdens the formal sector and encourages informality.[35] This vicious cycle is widely reported in the literature. Our modeling, discussed in Section 5, is directly based in the logic that the informal economy shifts the tax burden to the formal one.

Yet the linkages between informality and government reach beyond taxation issues and also interact with broader social impacts of informality, including social provision and inequality. Our modeling substantially represents the linkages from government revenues to social provision, because of the forward linkages in IFs from those revenues to expenditures. A government strapped for revenues by

[35] de Soto (1989: 155-157) points out, however, that the informal sector cannot avoid indirect taxes such as gasoline taxes or other sales taxes on inputs it purchases from the formal sector. Our modeling at this point treats indirect and direct taxes with the same logic of linkage to informality; in future revision we should revisit this and exempt indirect taxes from the assumption of non-payment in the informal sector.

informality will be much constrained in terms of spending of all sorts, including education, health, infrastructure, and R&D. Similarly, spending constraints reduce government transfer payments, which normally are at least intended to aid the poor and reduce inequality. Hence, we will also discuss here what the literature tells us about these secondary and broader impacts of informality on government and society.

The immediate loss of government revenue due to an increase in the informal sector is partially offset by the fact that informal workers generally do not receive the social security or health benefits that these revenues provide. Nevertheless, from a fiscal perspective, as Loayza (2005: 2) points out that, "the informal sector generates a negative externality that compounds its adverse effect on efficiency: informal activities use and congest public infrastructure without contributing the tax revenue to replenish it." This decreases the marginal value of public infrastructure for those in the formal sector who contribute to its maintenance, and closes the margin between the benefits they gain from staying in the formal sector and the taxes they pay for that status.

If enforcement mechanisms, which can take the form of government regulation or social norms, are respected, then countries should be easily able to support the cost of the informal sector; however, if incentives are strong enough for a noticeable number of actors to exit the formal sector, their non-compliance can signal serious long-term implications for the health of many social institutions. Centeno and Portes (2006: 38) describe this scenario, and its run-away consequences, in the following passage:

> *The informal economy frustrates the resolution of collective action problems or the adequate compensation for hidden costs that are normally solved by a modern state. By this we mean the kinds of costs not borne directly by the participants (be they employers or employees): long-term health, retirement, public safety, and social services. Informal activities also commonly carry "hidden" costs, such as environmental degradation or the appropriation of collective goods. Informal enterprise, from street vending to pirate production and sale of goods, frequently depends on the use of publicly provided services without contributing am thing to cover their costs.*
>
> *Because of the absence of resources and because of the pervasiveness of illegal activity, it is difficult for the state to sustain a credible role as universal enforcer of rules and contracts… Associated problems range from the massive corruption of poorly paid policemen to the unavailability of civil courts to settle contracts. This has enormous long-term consequences. Because the policeman or the judge cannot be trusted, citizens have no way of enforcing the most basic regulations… Civil society in such contexts finds its own mechanisms of self-regulation, but these do not provide a basis for sustained modern development.*

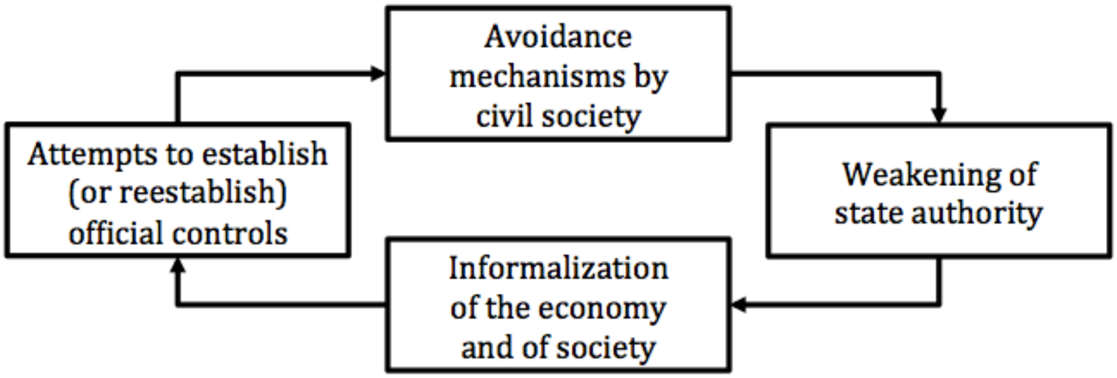

Figure 5.2. The process of informalization under "frustrated" states
Source: Adapted from Centeno and Portes (2006: 33)

The vicious cycle depicted in this narrative begins with an increase in non-compliance and a matched decrease in rule-of-law eroding the capabilities and legitimacy of public institutions. Decreasing revenue streams drive down investments in public goods. While the shift toward informality is due to the short-term benefits, mainly higher effective earnings, a continuation of this path undermines the public resources and institutions that are necessary to weather shocks and solve collective action problems in the long-run. This reduced capacity results in a segmentation of the economy, with growing inequality between those on the inside and those on the outside. Again Centeno and Portes (2006: 40) offer an eloquent narrative of this progression:

> *…the perpetuation of the informal economy also insures the continuation of vast social inequalities. A significant part of the economically active population laboring without protection and without the possibility of collective representation may comprise a considerable comparative advantage for some industries, but it bodes ill for the possibility of reducing inequality and wide-spread poverty in these nations. In a sense, the functionality of the informal sector for the state and firms in the formal economy depends precisely on the continuing vulnerability and poverty of those laboring in underground activity. This inequality, in turn, further limits the chances for the development of an effective state because those with resources to tax have the political weight to avoid such payments.*

Empirical support for this narrative can be found in Kim (2005), who notes a similarly positive relationship between informality and inequality across a large number of empirical studies, particularly ones focused on transition economies. Additionally, Chong and Gradstein (2004) create a model to test this hypothesis, and find evidence in support of the positive relationship between informality and inequality, as well as its connection to institutional quality.

The relationship between higher taxes, more informality, and lower institutional quality is, however, not inherently vicious. Friedman et al. (2000) find many countries that have both higher levels of taxes and lower levels of informality. Their explanation is that these countries are able to achieve such low levels of tax evasion, despite a higher tax burden and greater incentives to exit, because their societies perceive the regulations and enforcement mechanisms to be fair. Furthermore, complying with higher taxes means that productivity-increasing public goods receive adequate funding, which offers even more incentive for formal employment.

In summary, there are strong theoretical links between informality (or tax evasion), social provision, rule of law, and inequality. Under certain conditions they can create a vicious cycle that increases informality, decreases rule of law, and undermines the public intuitions that aid in the mitigation of collective needs and inequality. However, with public institutions and regulations that are perceived as fair by society, the story becomes one of increased social provision and productivity enhancing public infrastructure, which is able to support a healthier informal sector.

We turn now to a discussion of how we have modeled the informal sector and both its drivers and its forward linkages.

6. Structure of the Informal Economy Representation in IFs

Forecasting the informal economy

The approach we have taken to modeling the informal economy reflects the conceptualization and analytical work that previous sections described. Figure 6.1 provides a high-level schematic diagram of the informal sector in IFs. The central dynamics across time are generated in interaction with the broader IFs system. For the informal economy module we first use multiple driver variables within IFs to determine the share of the labor force that is informal. Additional equations split that informal labor share by destination (informal sector or formal and household sector) and by sex. The informal labor share and additional driver variables feed forward to the share of the official GDP that is informal (assuming that most countries, following the OECD's NOE convention as does Peru, estimate the size of the informal economy within official values of the GDP). The share of the economy that is informal has two direct forward effects, on multifactor or total productivity and on tax revenues; larger informal economy shares are a drag on both and thereby slow down GDP growth and constrain government finance. GDP and government finance are parts of the larger IFs forecasting system, thereby passing the impacts of changes in the informal economy size on to all other variables in IFs including the drivers of the informal labor share. This closed linkage pattern suggests that changes in the informal shares can generate a wide range of feedbacks, both positive (virtuous if the shares decrease or vicious if the shares increase) and negative (or equilibrating).

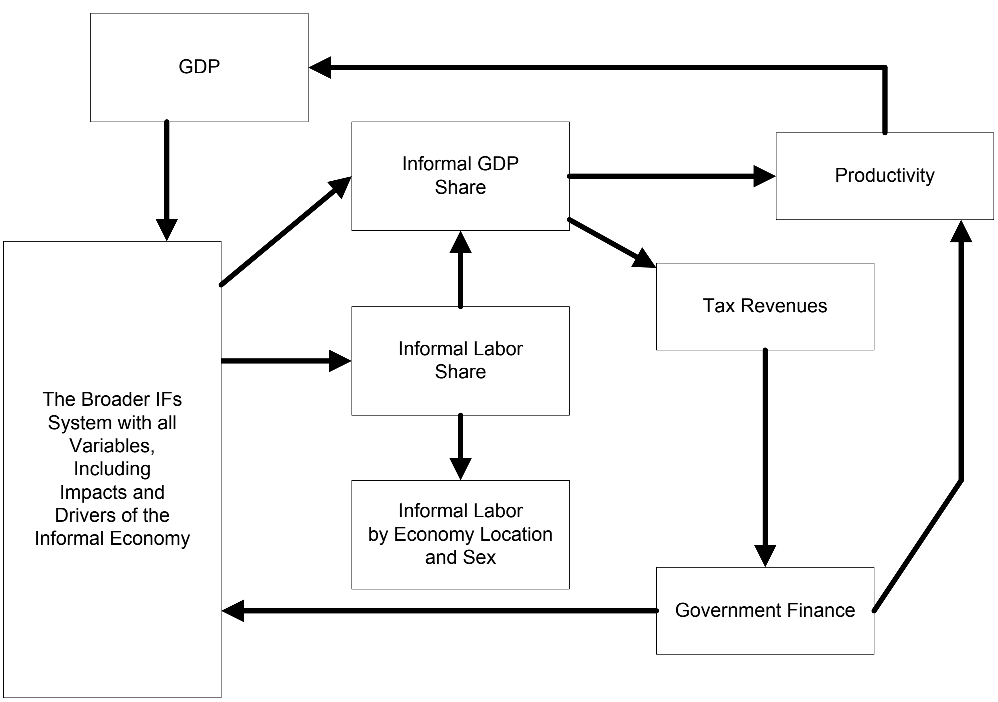

Figure 6.1. High-level overview of the forecasting in IFs of the informal economy
Source: The authors

Elaboration of core elements in forecasting the informal economy

Figure 6.2 elaborates the schematic diagram of Figure 6.1 with a more detailed picture of the core informal economy model. For instance, it shows that the drivers of the informal labor force size generally fall into three categories (see Section 4 of this report on drivers in the literature and in our empirical analysis). First are deep and slow-to-change socioeconomic variables. Although IFs uses GDP per capita at purchasing power parity across many of its models and submodels as a core driver, we determined that the years of formal educational attainment by adults (15 years of age and older) was a somewhat stronger variable in the analysis of the informal economy and has the additional virtue of suggesting some leverage points in the educational system for moving more of the labor force to the formal economy.

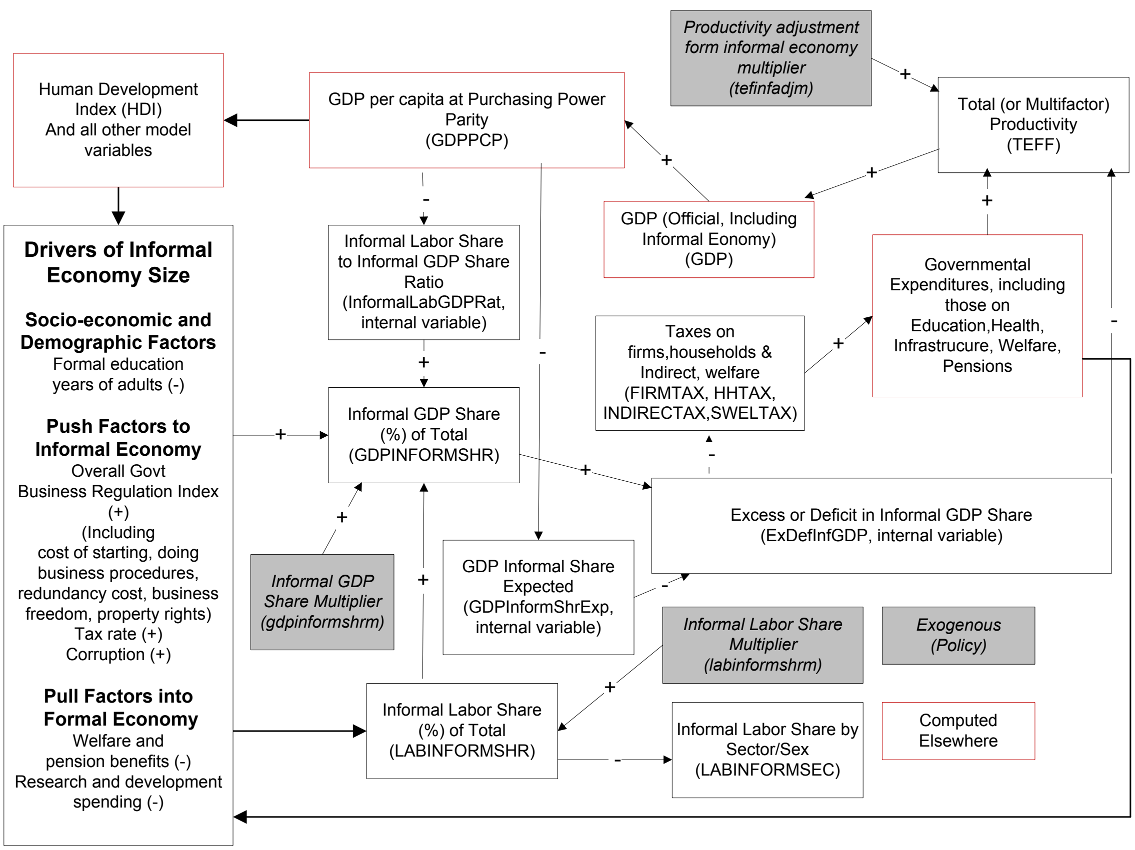

Figure 6.2. Core elaboration of the forecasting in IFs of the informal economy
Source: The authors

The next two categories are, respectively, push and pull factors that tend to keep (or push) people and businesses out of the formal sphere or pull them into it. On the push side are: (1) an overall government business index with the 5 components again identified and elaborated earlier, (2) the tax rate on corporate profits, and (3) the level of corruption. On the pull side are: (1) government spending on social welfare and pensions as a portion of GDP and the tax rate, which can be considered a proxy for more general strength of the governance capacity, and (2) spending on research and development as a share of GDP, which has an important independent, although relatively weakly supported, effect on informal GDP share.

In the upper right of the figure, there is some basic elaboration of the forward linkages of the informal GDP, and of formalizing it, to economic productivity and government revenues. Subsequent text will show still more detail and discuss the model formulations. Subsequent sub-sections will in turn elaborate the determination of the informal labor share (LABINFORMSHR), its split by location of the informal labor and sex (LABINSFORMPCNTINF and LABINFORMPCTNONINF), the determination of the informal GDP share (GDPINFORMSHR), and the forward linkages of informal GDP share to productivity and government revenues.

Forecasting informal labor share

We described our estimation of the formulations that drive the informal labor share in Section 4. The value of Figure 6.3 lies in its further identification of the variables in the larger IFs system actually used in the equations, for instance for years of formal adult education (EDYRSAG15). See the documentation on the various IFs models or the Scenario Manual for insight with respect to the parameters that can manipulate these four drivers.

In our building of the model of informal economy we paid particular attention to education of adults as a long-term or deep driver (as is GDP per capita, with which it correlates very highly). The importance of the education variable can be seen in the r-squared of 0.29 between it and informal labor share (see again Figure 4.3).

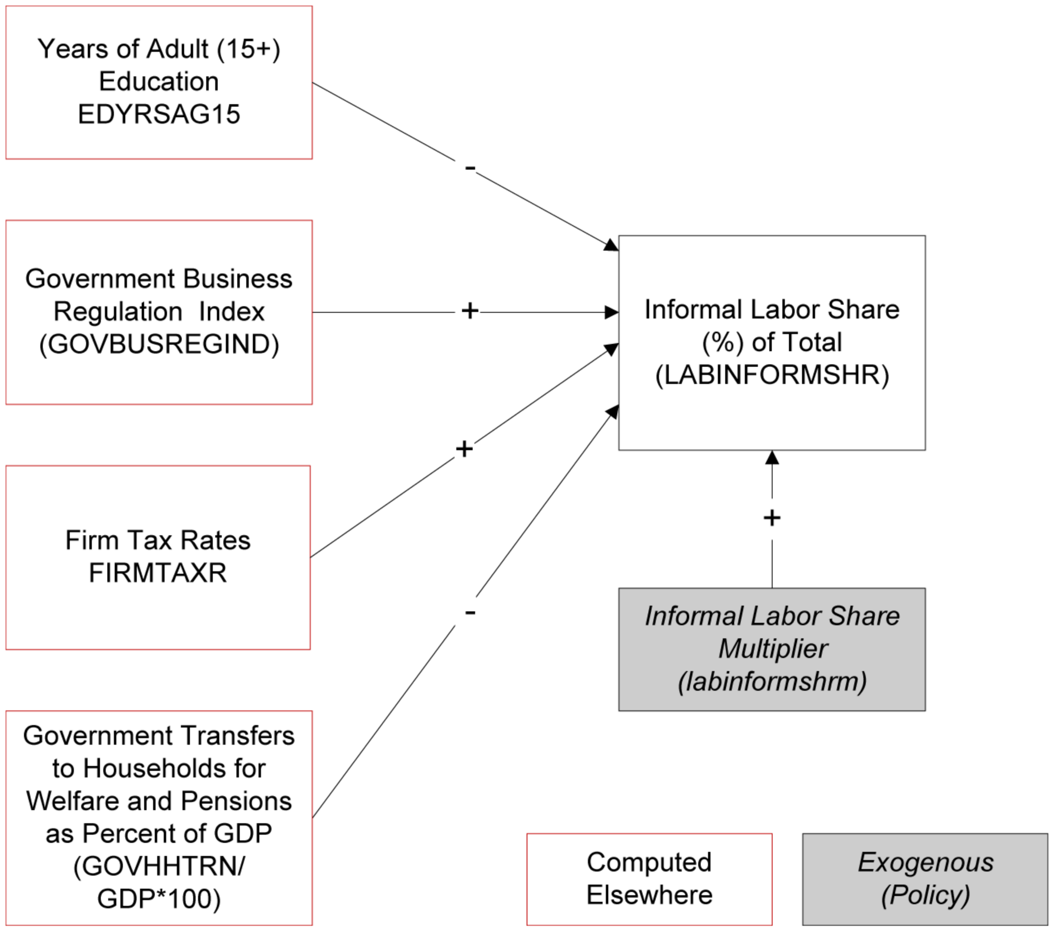

Figure 6.3. Forecasting informal labor share
Source: The authors

In early work we explored keeping that relationship with adult education years as a fixed core in forecasting informal labor share—using other variables to explain only the residual variation it leaves in informal labor share. We moved away from that approach in part because the other drivers are both very important and are themselves sometimes long-term and deeper drivers reflecting socioeconomic change.

The forecast of informal labor share as a percentage of total labor force (LABINFORMSHR) is computed in two primary steps. First, a basic value of that percentage (LabInformal) is calculated as a function of the four drivers identified in Figure 6.3. The driver terms are all from the previous model year (t-1), because in the recursive structure of IFs they are computed later in the sequence of variables.

$$
\begin{gathered}
LabInformal_{r,t} = \boldsymbol{labinformcoeffintercept} \\
+\boldsymbol{labinformcoeffed} * EDYRSAG15_{r,t-1,s=3} \\
+\boldsymbol{labinformcoeffbus} * GOVBUSREGIND_{r,t-1} \\
+\boldsymbol{labinformcoefftax} * \mathrm{Log}(\mathrm{FIRMTAXR}_{\mathrm{r,t-1}})
\end{gathered}
$$

$$+\boldsymbol{labinformcoeffhhtrn} * \left(100 * \frac{GOVHHTRN_{r,t-1}}{GDP_{r,t-1}}\right)$$

In the above equation, the values of the parameters are: ***labinformcoeffintercept***=-12.613***, labinformcoeffed=***-2.510***, labinformcoeffbus =***12.604***, labinformcoeffhhtrn=***-1.810***, labinformcoefftax=***10.6. And in the above equation, education years are for the total population (sex=3) rather than males or females.

The second step modifies that initial calculation by converging a base year shift term (LabInformShrShift) to zero over years specified by labinfconvyr and by applying an exogenous multiplier (***labinformshrm).***

$$\begin{aligned}&LABINFORMSHR_{r,t}\\&= (LabInformal_{r,t}\\&+ ConvergeOverTime\left(LabInformShrShift_{r,t=1}, 0, labinfconvyr_{r,t}\right))\\&* \boldsymbol{labinformshrm}_{r,t}\end{aligned}$$

where

$$LabInformShrShift_{r,t=1} = LABINFORMSHR_{r,t=1} - LabInformal_{r,t=1}$$

and

$$labinfconvyr_{r,t_r} = 60 \text{ if } LabInformShrShift_{r,t=1} > 0$$
$$labinfconvyr_{r,t_r} = 120 \text{ if } LabInformShrShift_{r,t=1} \leq 0$$

In the above formulation note that an adjustment to the calculation of LABINFORMSHR is required to smooth the transition from historical data to forecast values, because the initial (t=1) value of that variable from data will not be consistent with the calculation of it in the function that generates LabInformal. The adjustment or shift factor (LabInformShrShift) is the difference in the first year between the value of LABINFORMSHR and the calculated value from the driver function. The shift factor constrains the model output in the initial year to better match historical data and then gradually lets the forecasting formulation take over. It would be possible to maintain that additive shift factor as a constant over time, but because both informal labor shares generally decline with economic growth, we converge that factor over time toward 0 or lack of adjustment. If the shift factor is greater than 0, we converge it in 60 years; otherwise in 120 years. The reason for the slower upward convergence is that, in general, the informal labor share decreases with development and advance in GDP per capita; upward movement would be very unusual.[36]

[36] Convergence parameters are added to the model when data values for the base year (2010) of variables differ substantially from those that forecasting functions also provide for that base year. In this situation we compute shift factors that record the discrepancy in the base year. Although those shift factors could be maintained as constant adjustments to forecasting formulations over time, our experience is that the reasons for the discrepancies fall into 2 categories: (1) errors in data that we have no ability to eliminate and (2) influences of variables not in our forecasting formulation, such as country-specific historical path dependencies related to historical events or cultural patterns. Over time data errors tend to be corrected and

Forecasting informal labor share by sector

It is very useful in our forecasting to know more about the makeup of informal labor. In particular, we want to know more about where that labor is employed and how it breaks down by sex. With respect to where it is employed, we use the categorization discussed in Section 2 and often used by ILO-WEIGO, namely the portion of informal labor employed outside the informal sector (in the formal sector or in households) and the portion employed informally. We further divide each by sex.

The first step in computation is of the portion of informal labor employed inside the informal sector (LABINFORMPCNTINF), which varies inversely with the total share of labor that is informal (LABINFORMSHR). That is, as the informal share of total labor rises, a smaller portion of the total informal labor will be able to find positions in the formal sector (or as the informal share of total labor falls, the portion of the remaining informal labor employed in the formal sector will rise). This is computed for both sexes (s=Total) prior to a subsequent division by sex.

$$LABINFORMPCNTINF_{r,s=Total,t} = \boldsymbol{infinformcoeffintercept} + \boldsymbol{infinformcoefflabinf} * LABINFORMSHR_{r,t})$$

The values of the parameters currently are: ***infinformcoeffintercept***=55.243; ***infinformcoefflabinf***=- 0.415.

The percentage of the informal labor that is outside the informal sector (LABINFORMPCNTNONINF) is simply the residual.

$$LABINFORMPCNTNONINF_{r,s=Total,t} = 100 - LABINFORMPCNTINF_{r,s=Total,t}$$

The further division by sex of the informal labor by sector is done by using the initial patterns by sex (LABINFORMPCNTINFFemShare and LABINFORMPCNTNONINFFemShare). There are not enough data and conceptual foundation to establish a dynamic relationship for the share by sex.

---

country-specific influences tend to erode, suggesting that we should generally converge the shift factors toward 0 over time. Because both bases for the shift factor involve information that we cannot have, the convergence time is a parameter that must be set subjectively. Our basis for doing so is analysis of model behavior with longer or shorter shift factors, with an eye especially to avoiding transients in behavior by having convergence times that are too short. Thus most convergence times are set at 50 years or more, with such attention to their implications for model behavior. In cases where we believe that shifts are most likely to occur in one direction (e.g. we expect more movement downward in shares of economies that are informal than upward), we will set the convergence times differentially as in the equation indicated.

$$LABINFORMPCNTINF_{r,s=Female,t} = LABINFORMPCNTINF_{r,s=Total,t} * LABINFORMPCNTINFFemShare_{r,t=1})$$

$$LABINFORMPCNTINF_{r,s=Male,t} = LABINFORMPCNTINF_{r,s=Total,t} - LABINFORMPCNTINF_{r,s=Female,t}$$

where

$$LABINFORMPCNTINFFemShare_{r,t=1} = \frac{LABINFORMPCNTINF_{r,s=Female,t=1}}{LABINFORMPCNTINF_{r,s=Total,t=1}}$$

Similarly,

$$LABINFORMPCNTNONINF_{r,s=Female,t} = LABINFORMPCNTNONINF_{r,s=Total,t} * LABINFORMPCNTNONINFFemShare_{r,t=1})$$

$$LABINFORMPCNTNONINF_{r,s=Male,t} = LABINFORMPCNTNONINF_{r,s=Total,t} - LABINFORMPCNTNONINF_{r,s=Female,t}$$

where

$$LABINFORMPCNTNONINFFemShare_{r,t=1} = \frac{LABINFORMPCNTNONINF_{r,s=Female,t=1}LABINFORMPCNTNONINFFemShare_{r,t=1}}{LABINFORMPCNTNONINF_{r,s=Total,t=1}}$$

Forecasting informal GDP share

In Section 4, we described our estimation of the formulations that drive the informal GDP share. Figure 6.4 identifies again the variables used in the model, namely the informal labor share (LABINFORMSHR), the level of corruption (GOVCORRUPT), and the expenditure as a percentage of GDP by the society on research and development (RANDEXP).

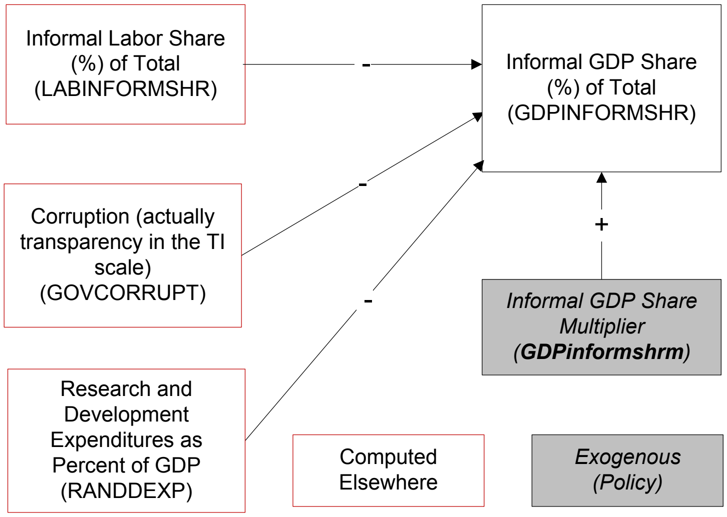

Figure 6.4 Forecasting informal GDP share
Source: The authors

As in the computation of informal labor share, the forecast of informal GDP share as a percentage of GDP (GDPINFORMSHR) involves two major steps. First, a basic value of that percentage (GDPInformal) is calculated as a function of the four drivers identified in Figure 6.4. The driver terms are all from the previous model year (t-1), because in the recursive structure of IFs they are computed later in the sequence of variables.

$$\begin{gathered} GDPInformal_{r,t} = \boldsymbol{gdpinformcoeffintercept} \\ +\boldsymbol{gdpinformcoefflabinf} * LABINFORMSHR_{r,t-1} \\ +\boldsymbol{gdpinformcoeffcorrupt} * GOVCORUPT_{r,t-1} \\ +\boldsymbol{gdpinformcoeffRandD} * RANDDEXP_{r,t-1} \end{gathered}$$

In the above equation, the values of the parameters are: ***gdpinformcoeffintercept*** = 18.861***, gdpinformcoefflabinf =*** 0.193***, gdpinformcoeffcorrupt =*** -1.585***,*** and ***gdpinformcoeffRanD =*** -2.6. The negative value on the corruption term is potentially misleading because in Transparency International's measure, higher values are more transparent and lower ones are more corrupt; the equation term increases GDPInformal as corruption increases.

The second step modifies that initial calculation by converging a base year shift term (GDPInformShrShift) to zero over a number of years specified by gdpinfconvyr and by applying an exogenous multiplier (***gdpinformshrm).***

$$\begin{aligned} & GDPINFORMSHR_{r,t} \\ & = (GDPInformal_{r,t} \\ & + ConvergeOverTime\left(GDPInformShrShift_{r,t=1}, 0, gdpinfconvyr_{r,t}\right)) \\ & * \boldsymbol{gdpinformshrm}_{r,t} \end{aligned}$$

where

$$GDPInformShrShift_{r,t=1} = GDPINFORMSHR_{r,t=1} - GDPInformal_{r,t=1}$$

and

$$gdpinfconvyr_{r,t_{r}} = 70 \text{ if } GDPInformShrShift_{r,t=1} > 0$$
$$gdpinfconvyr_{r,t_{r}} = 200 \text{ if } GDPInformShrShift_{r,t=1} \leq 0$$

In the above formulation note that an adjustment is required to the calculation of GDPINFORMSHR, because the initial (t=1) value of that variable from data will not be consistent with the calculation of it in the function that generates GDPInformal. The adjustment or shift factor (GDPInformShrShift) is the difference in the first year between the value of GDPINFORMSHR and the calculated value from the driver function. It would be possible to maintain that additive shift factor as a constant over time, but because both informal GDP shares generally decline with economic growth, we converge that factor over time toward 0 or lack of adjustment. If the shift factor is greater than 0, we converge it in 70 years; otherwise in 200 years. The reason for the slower upward convergence is that, in general, the informal labor share decreases with development and advance in GDP per capita; upward movement would be very unusual.

Shadow Economy

Shadow economy as a share of total is calculated using a formulation similar to that used for the informal share of the economy. Estimations from Schneider and Elgin were used to initialize the model in the base year (countries with missing data are initialized using a regression of the most recent year data against per capita income of countries). The regression model for shadow economy uses the same set of independent variables as did the model for the informal economy, i.e., informal labor share (LABINFORMSHR), corruption (GOVCORRUPT) and total R&D spending (RANDDEXP).

$$\begin{aligned} GDPInformalCalc_{r,t} &= \boldsymbol{gdpshadowcoeffintercept} + \boldsymbol{gdpshadowcoefflabinf} \\ &* LABINFORMSHR_{r,t-1} + \boldsymbol{gdpshadowcoeffcorrupt} \\ &* GOVCORUPT_{r,t-1} + +\boldsymbol{gdpshadowcoeffRandD} * RANDDEXP_{r,t-1} \end{aligned}$$

where, r and t denote country/region and time, respectively,

In the above equation, the basic values of the parameters are: ***gdpshadowcoeffintercept*** = 36.2814***, gdpshadowcoefflabinf =*** 0.1019***, gdpshadowcoeffcorrupt =*** -1.8849***,*** and ***gdpshadowcoeffRanD =*** -2.9572. The negative value on the corruption term is potentially misleading because in Transparency International's measure, higher values are more transparent and lower ones are more corrupt; the equation term increases GDPInformal as corruption increases.

As discussed above for the informal economy, initial values and base year model calculations for the size of the shadow economy will typically differ. Shift factors are used to adjust computed values to those from data in the base year, and the shift factors converge to zero over time.

For countries where the data on GDP share of shadow economy is greater than the informal labor share, indicating a somewhat counterintuitive situation in which the shadow sector seems to have a higher productivity per person than the formal economy, we accept those values in the model base year, but converge shadow GDP share to the informal labor shares over a period of 25 years.

Forward linkages: general approach

Figure 6.5 elaborates the earlier core portrayal of forward linkages in Figure 6.2. The fundamental driver of changes in economic productivity and government finance is the excess or deficit in the informal GDP share (ExDefInfGDP), relative to the expected value (see the explanation earlier).

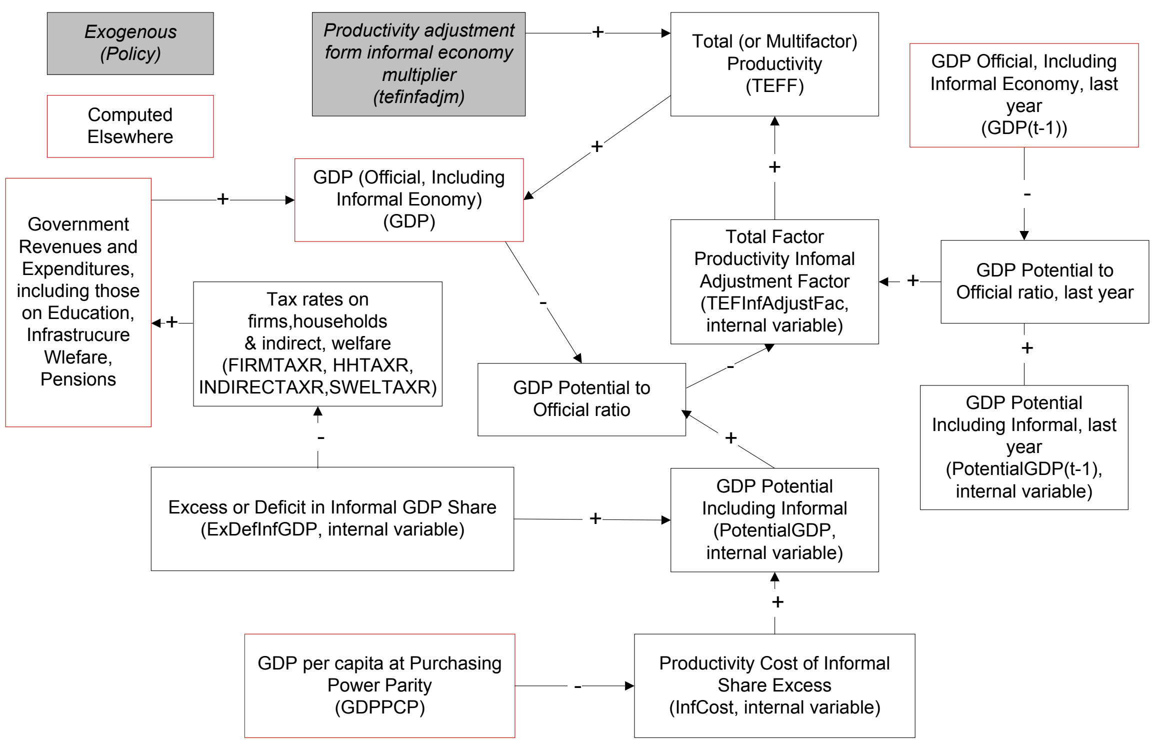

Figure 6.5. Forward linkages of informal share size to productivity and revenues
Source: The authors

We begin by elaborating the manner in which that excess or deficit affects total (or multifactor) productivity and then turn to the manner in which it affects tax rates and therefore government revenues and expenditures. Reference to Figure 6.5 as we explain the equations below may be helpful.

Forward linkages: productivity and GDP

We have structured the forward linkages so that the model user can drive them with either the size of the informal economy (GDPINFORMSHR) or the size of the shadow economy (GDPSHADOWSHR). The logic is the same in either case. In the discussion below, which is used will depend on the value of the shadow economy switch (gdpshadowon). If that parameter is 0, as it is in the model base case, the driver is the informal economy share; if it is 1, the driver is the shadow economy share.

The key to the productivity linkage is to recognize that an informal economy (or a shadow economy but to a much lesser degree) has the tendency, particularly in developing countries, to generate an official GDP that is lower than the potential GDP (PotentialGDP), because labor is employed with less capital and less efficiency in the informal sector. The magnitude of this productivity cost of informality (InfCost) can be illustrated by the fact that the data for some economies suggest that while 60 percent or more of labor can be informal, only 20 percent of GDP might be informal. This suggests that each unit of informal labor in such a country is used only about one-third as productively as is formal labor (we compute the ratio as the inverse, suggesting that each surplus unit of informal or shadow economy in this example could have a productivity cost of 3; we further bound that ratio so that it cannot exceed 3 or fall below 1).

If gdpshadowon=0 then

$$InfCost_{r,t} = \frac{LABINFORMSHR_{r,t}}{GDPINFORMSHR_{r,t}}$$

else

$$InfCost_{r,t} = \frac{LABINFORMSHR_{r,t}}{GDPSHADOWSHR_{r,t}}$$

It is important to emphasize, however, that the productivity and GDP growth rates in the basic IFs model already reflect the size of each country's informal GDP share and whatever drag or boost that share is giving to growth. Moreover, normal patterns of change in the informal GDP share (and we have seen that the share tends to decrease over the long-run with either advance in adult formal education or GDP per capita, are essentially already built into the long-term dynamics of productivity change in IFs, for instance the basic convergence or catch-up of productivity in developing countries with higher income countries. Thus, our attention here is not to the basic impact of informality on growth, but the incremental impact that deviations from the normal or expected size of the informal sector and the normal or expected change in that size might have. We can refer to such deviation as excess or deficit informal GDP share (ExDefInfGDP).

This point is so important that it bears rephrasing and repeating in only slightly different language. Why is this excess or deficit important? The dynamics of the IFs system's basic forecasting of economic productivity and GDP and of government revenue and expenditure patterns are all built on historical patterns and estimated functions that represent the normal processes within countries of gradual reduction of informal economy size with socioeconomic development. Thus, a *normal* decrease over time in the informal economy size should not be allowed to further affect the dynamics of productivity or government finance. Similarly, if a country has a persistent and stably larger or smaller informal economy than would be expected at its level of socio-economic development, with GDP per capita at PPP serving as a proxy for that—it should have normal productivity and finance dynamics. It will be countries that change the share of the informal economy relative to their expected values with development, and that therefore demonstrate changes in the excess or deficit of the informal GDP share (ExDefInfGDP) that will alter their expected development pattern.

The discussion above detailed the calculation of the forecast value of the actual informal share of GDP (GDPINFORMSHR), so to compute the excess or deficit informal share we need only to explain the calculation of the expected share (GDPInformShrExp). That calculation uses a function estimated against GDP per capita at PPP, namely "GDP/Capita (PPP 2011) Versus Informal Share of GDP (Most Recent) Log."

$$GDPInformShrExp_{r,t} = F\left(GDPPCP_{r,t}\right)$$

$$GDPShadowShrExp_{r,t} = F\left(GDPPCP_{r,t}\right)$$

This expected level is then subtracted from the actual forecast of the informal share, and the difference is divided by 100 to calculate the excess or deficit (ExDefInfGDP) in the extent of informal activities in ratio rather than percentage terms.

If gdpshadowon=0 then

$$ExDefInfGDP_{r,t} = (GDPINFORMSHR_{r,t} - \; GDPInformShrExp_{r,t})/100$$

else

$$ExDefInfGDP_{r,t} = (GDPSHADOWSHR_{r,t} - \; GDPShadowShrExp_{r,t})/100$$

What might be the potential GDP of an economy with an informal percentage share that is larger than expected? We can calculate that by multiplying the Official Economy (GDP) by the excess or deficit share (ExDefInfGDP) and by the unit informal cost (InfCost), adding that to base share of the official economy—the base share is the official economy minus that excess or deficit informal portion and should not be understood to be entirely formal, but rather to be the expected core mix of formal and informal.

$$PotentialGDP_{r,t} = OfficialGDP_{r,t} * (1 - ExDefInfGDP_{r,t}) + OfficialGDP_{r,t} * ExDefInfGDP_{r,t} * InfCost_{r,t}$$

To illustrate this more concretely, if a country had an official economy of $100 billion, but 20 percent of that were "excess" informal economy, the base economy would be $80 billion. If the unit cost of excess informality were a factor of 3, the excess $20 billion of informal economy could potentially become $60 billion and boost the potential GDP to $140 billion. Thus the calculation of the potential GDP in the formula above would be 100*(1-0.2) + 100* 0.2*3 = 80+60=140.

For the calculation further below of impact on productivity of reducing (or increasing) this potential GDP, we need a ratio of the official to potential GDP and will need to look at that ratio over time. Getting that ratio only requires a re-ordering of the terms in the above equation:

$$\frac{OfficialGDP_{r,t}}{PotentialGDP_{r,t}} = \frac{1}{(1 - ExDefInfGDP_{r,t}) + ExDefInfGDP_{r,t} * InfCost_{r,t}}$$

We call the ratio of official to potential GDP the (potential) adjustment factor for the total factor productivity and thus economy (TefInfAdjust).

$$TefInfAdjust_{r,t} = \frac{OfficialGDP_{r,t}}{PotentialGDP_{r,t}}$$

Again, we can clarify by extending our concrete illustration. For the economy above the right hand side is 1/((1-0.2)+0.2*3)=1/(0.8+0.6)=1/1.4=0.714. Assume that the country were able to reduce its excess informal economy to 10 percent. The equation would then become 1/((1-0.1)+0.1*3)=1/(0.9+.3)=1/1.2=0.833.

If the country accomplished this reduction (which might take 10 or 20 years to do, of course), what would be the impact on economic production? Basically, it should be able to boost the overall magnitude of production by 0.833/0.714=1.167, that is, by 17 percent.

The change in the official/potential ratio from year-to-year (TefInfAdjustFac) is what will affect productivity and economic growth. The model user can also modify the impact with an adjustment multiplier parameter (***tefinfadjm***).

$$TefInfAdjustFac_{r,t} = \frac{TefInfAdjust_{t}}{TefInfAdjust_{t-1}} * \boldsymbol{tefinfadjm_{r,t}}$$

We then feed that change into a variable that the model maintains for total factor productivity (TEFF). That variable is a stock, so that it carries over from year to year and has been adjusted already by other drivers of productivity before it is adjusted below by the informal economy term. Thus, our example country making a 10 percent reduction in excess informal GDP over 10 years might actually make a 1.6 percent change for each of the 10 years. Those annual changes would cumulatively affect TEFF in such a way that after 10 years, the total impact should be roughly that of the 17 percent computed above (roughly, because compounding and internal model dynamics would change that somewhat). TEFF is maintained by economic sector (s).

$$TEFF_{r,t,s} = TEFF_{r,t,s} * (1 + TefInfAdjustFac_{r,t})$$

The variable TEFF is a multiplicative term in the production function and the computation of value added and GDP. Thus, any change in TEFF feeds through to GDP in the same proportion.

Forward linkages: government revenue

The general logic for computing the impact on productivity and GDP from changes in excess or deficit informal GDP share is also used for computing its impact on government revenues, except that there is no need for a unit cost term. Except for external items such as foreign aid, government revenues in IFs come from four taxes: those on firm profits (FIRMTAX) and on household income (HHTAX), plus specialized taxes such as indirect taxes on consumption (INDIRECTTAX) and contributions to social welfare programs or pensions (SSWELTAX). In each case, IFs also has a variable that specifies the rate of the tax on the appropriate base (FIRMTAXR, HHTAXR, INDIRECTTAXR, and SSWELTAXR). See again Figure 6.2.

The simplifying assumption we make is that informal economies contribute very little or nothing to the true base for these taxes. Thus the tax rates (coming to us along with tax bases and amounts from official data) are actually in the real world associated with the formal sub-portion of the official economy and appropriate tax bases in that official economy. As the share of the informal economy decreases, we can expect those bases to expand and to generate effectively higher tax rates. In IFs, we can use this general understanding to increase (or decrease) the tax rates accordingly as the excess or deficit informal economy variable decreases (or increases).[37] We can compute a tax rate adjustment factor (TaxInfAdjustmentFac) that reflects this, also introducing a multiplicative parameter usable for scenario analyses (***taxinfadjm***).

[37] The model computes HHTax=HHIncOfficial*HHTaxR, but the actual situation we are trying to represent is HHTax = HHIncFormal* (HHTaxR * X) where X is an implicit factor of increase in the real tax rate so as to represent the application of it only to the formal economy. Some algebra would change the necessary adjustment in the tax rate, but it should not be highly significant relative to other sources of uncertainty.

$$TaxInfAdjustmentFac_{r,t} = 1 + \left(ExDefInfGDP_{r,t=1} - ExDefInfGDP_{r,t}\right) * \boldsymbol{taxinfadjm_{r,t}}$$

Because the adjustment factor is computed relative to the first year (t=1), the change in all four tax rates should also be applied to the rate of the first year. (In the base of households there are two types (h), skilled and unskilled.)

$$FIRMTAXR_{r,t} = FIRMTAXR_{r,t=1} * \ TaxInfAdjustmentFac_{r,t}$$

$$HHTAXR_{r,h,t} = HHTAXR_{r,,h,t=1} * \ TaxInfAdjustmentFac_{r,t}$$

$$INDIRECTTAXR_{r,t} = INDIRECTTAXR_{r,t=1} * \ TaxInfAdjustmentFac_{r,t}$$

$$SSWELTAXR_{r,t} = SSWELTAXR_{r,t=1} * \ TaxInfAdjustmentFac_{r,t}$$

Total government revenues are responsive primarily to these tax rates and their bases and the model includes dynamics that also link government expenditures to taxes and fiscal balances. Thus changes in these revenues will have pervasive effects through the rest of the model.

We can explore the actual impact of changes in all driver variables on both economic growth and government revenues when we turn to the use of the model in the next section.

7. Analysis

The analysis here, focusing entirely on Peru through 2050, uses the new representations of the informal economy that the previous section described. It provides a foundational exploration of that sub-model and its integration with the International Futures (IFs) system. The analysis has multiple elements. First, we look at the Base Case forecast, considering the current and prospective size of the Peruvian informal economy as represented by both the informal economy and shadow shares of GDP and the informal labor share of total labor.[38]

Then we turn to scenario analysis. The first alternative scenario explores forward linkages and has the character of a counterfactual analysis: what would happen if the decline in informal GDP share were to be considerably more rapid than in the Base Case? Specifically, what would be the impacts on economic productivity and government revenues?

The analysis then turns to sensitivity analysis of backward linkages, or drivers. It explores the implications of aggressive change (presumably via governmental policy action) in each of the drivers of informal labor and GDP shares individually.

Finally, we combine changes in all the drivers to explore a best case scenario with respect to the reduction of the informal economy. To signal the result of that exploration, we find that this combined intervention or best case scenario has considerably greater impact on the potential GDP of Peru in 2050 than does even the aggressive counterfactual analysis. Why would that be so? Because the actions that would reduce the size of the informal sector, such as advancing education and research and development or reducing corruption, not only reduce the size of the informal economy, they also have broader and reinforcing consequences in the IFs system (as they would in any country that pursued them). In short, there would be a multiplicity of positive implications associated with such an action plan.

Base Case analysis

Peru's economy has one of the highest informality profiles in the world. Informal labor and GDP shares, as shown in Figure 7.1, are high even relative to the substantial levels of most Latin American countries. Yet the historical data (the blended series discussed earlier in this paper and consistent for Peru with data from INEI), have shown some decline. And by mid-century, our forecasted declines are significant, cutting both shares by more than half. Even then, the informal labor share could be approximately 27 percent of the labor force, more than three times

[38] As indicated in earlier sections of this report, we have very considerable skepticism about the database for the shadow economy, which is largely tied to model-based estimates in the literature rather than to survey sources as is the database for the informal economy. We include forecasts related to it in the analysis here in response to requests from our CEPLAN colleagues for its representation in the model.

that of most contemporary high-income countries, and the shadow economy size could be nearly as large.

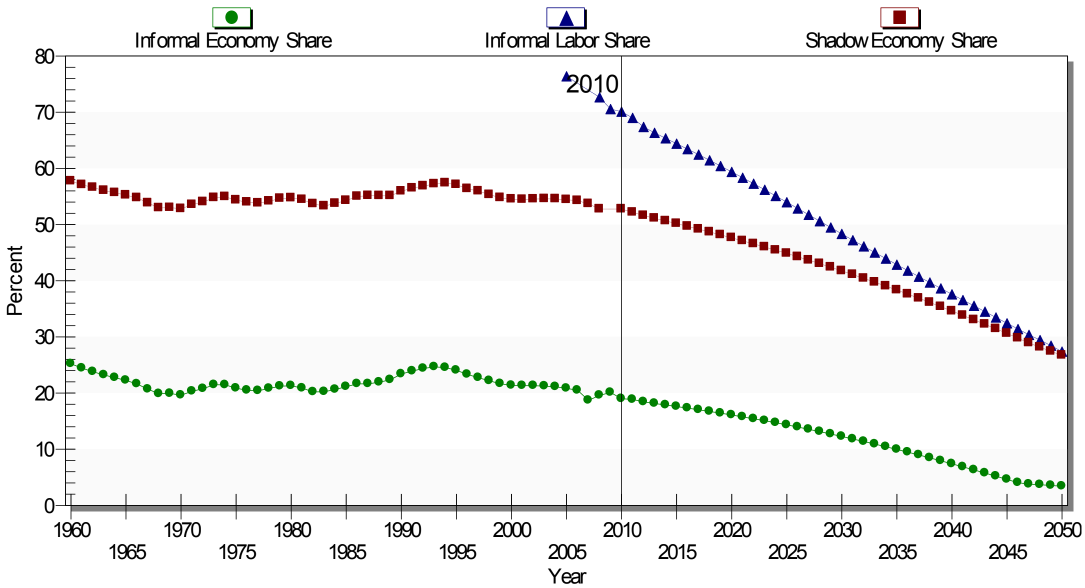

Figure 7.1. Informal labor and GDP shares for Peru, Base Case scenario, 2010—2050 and history
Source: IFs version 7.14

One of the issues to which we have given considerable thought in the construction of the model and our analysis of its forecasts is the likely pattern of change in countries around the world over the coming decades. As Perry et al. (2007) discussed, not only was there no significant decline in informal shares in most Latin American countries in the 1990s, there was increase in many. The reasons for that are complex, in part reflecting such regional and even country-specific factors as the so-called lost decade of economic growth in the 1980s, the playing out of the last stages in long-running demographic booms, inadequate educational systems, persistently high levels of socioeconomic inequality, and unfriendly policies of many governments toward formal businesses. Yet, in the context of the pronounced global relationship between those informal sector shares and the advance of deep development drivers such as education and GDP per capita (see again Section 4), it seems highly probable that the region's and Peru's future will see significant decline of informality with ongoing development processes.

The recent and likely-to-be continued decline in informal labor share does not mean, however, that there will be an immediate numerical decline in the total number of informal workers in Peru. Figure 7.2 shows that, because of the relatively youthful demographic structure of Peru and the continued growth of the overall labor force, the total number of informal workers in Peru is likely to remain just above 9 million into the 2020s, preceding a fairly steady and long-term overall decline. Our data suggest that females constitute a majority of the total informal labor force of the country. The forecast for continuation of that situation should be treated very cautiously, however, because we have not had the statistical foundation to create a dynamic forecasting representation of sex balance in the informal work force.

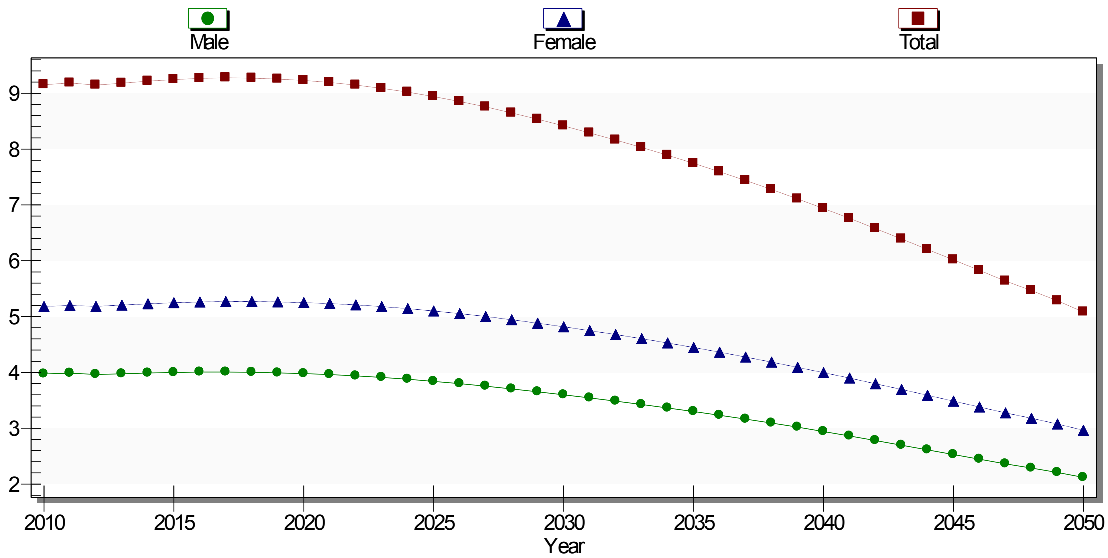

Figure 7.2. Informal labor in Peru by sex, Base Case scenario, 2010—2050
Source: IFs version 7.14

Another change that is highly likely in Peru is that a decreasing share of the informal labor will be in the strictly informal sector. As Figure 7.3 suggests, by 2050 there will be nearly as high a portion of informal labor outside of that sector, namely in the formal sector and in household labor as within.

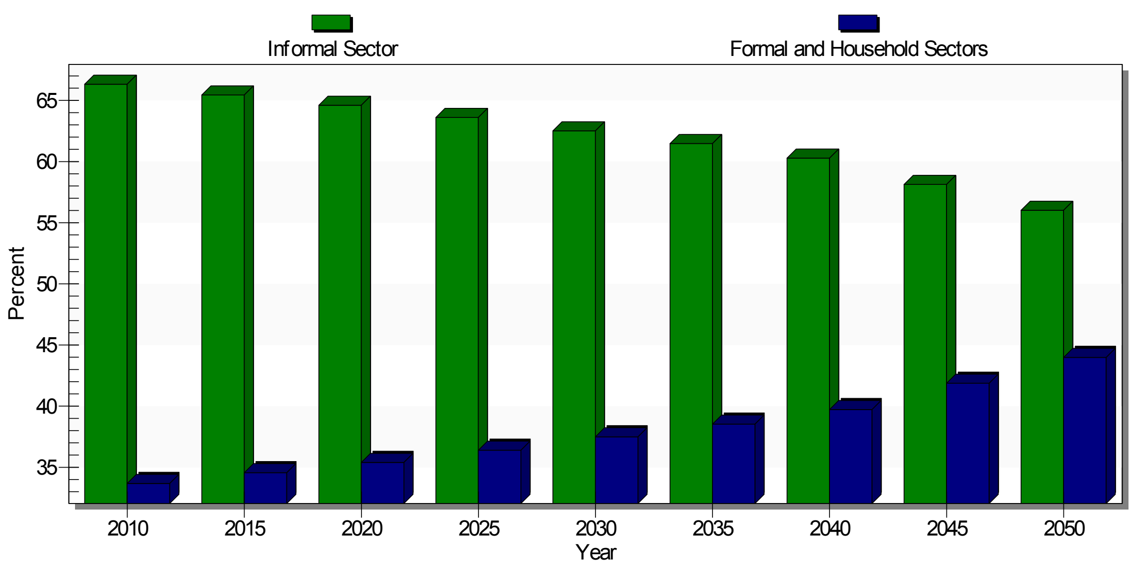

Figure 7.3. Informal labor in Peru by location in economy, Base Case scenario, 2010—2050
Source: IFs version 7.14

The model includes parameters allowing the user to adjust the Base Case pattern of change (see the model discussion in the previous Section), and it may be that further analysis would build an adjusted Base Case with slower decline than that in Figure 7.1. In fact, some analysis might well build a counterfactual forecast with little or no informal share decline. Here, however, we see reason to believe that such decline will occur, and we continue to build on the Base Case as now structured. In fact, the counterfactual scenario to which we now turn involves acceleration of the decline seen in Figure 7.1.

Counterfactual: decreasing the informal GDP share and examining impact

The purpose of the Counterfactual scenario is to isolate the relationship between the informal share of GDP and the variables that the share affects in the IFs model. To do this, we changed a single parameter for Peru, the multiplier on informal GDP share (***gdpinformshrm***), reducing it from 1 to 0.05 in the 15 years between 2016 and 2030—in essence simulating the elimination of informality. Figure 7.4 shows the impact of the intervention on the informal share, which it reduces to the minimum 1.0 percent that the model allows.

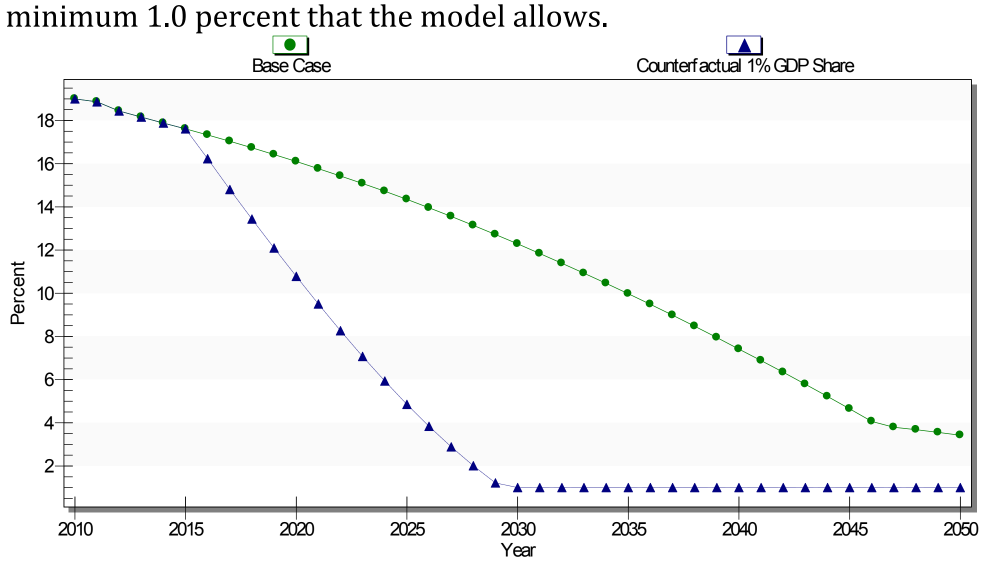

Figure 7.4. Informal GDP share of Peru in Base Case and Counterfactual scenarios, 2010—2050

Note: Although not shown, we built a similar counterfactual scenario for the shadow economy, using the parameter gdpshadowshrm. It similarly forces to size of the shadow economy to near zero by 2030.
Source: IFs version 7.14

The two variables of most immediate impact from such a model intervention, economic productivity and government revenues, have larger implications. Total or multifactor productivity enters the Cobb-Douglas production function of IFs as a multiplier on capital and labor contributions. A change in productivity thus directly and proportionately affects GDP. Figure 7.5 shows the resultant impact on GDP of our intervention (and Figure 7.6 shows the impact of the same counterfactual scenario applied to the shadow economy).[39] At its peak near the end of the of intervention period, it increases GDP by more than 26 percent relative to the Base Case. This gain erodes somewhat to about 15 percent by 2050 because the informal economy share would have been lower by then even in the Base Case (as we saw in Figure 7.4)

The other effect, however, is to government revenues and the taxes imposed on the economy, which also rise sharply with the posited conversion of informal economy into formal economy. By the end of the phase-in period in 2030 the scenario has increased government revenues by more than 2 percent of GDP and the premium relative to the Base Case again continues to rise (to more than 10% by 2050). The use of these revenues in the model includes investment in productive expenditures such as education, health, and infrastructure, as well as potential reduction in other taxes and increases in potentially less productive spending. There is, therefore, some secondary benefit for economic growth also.

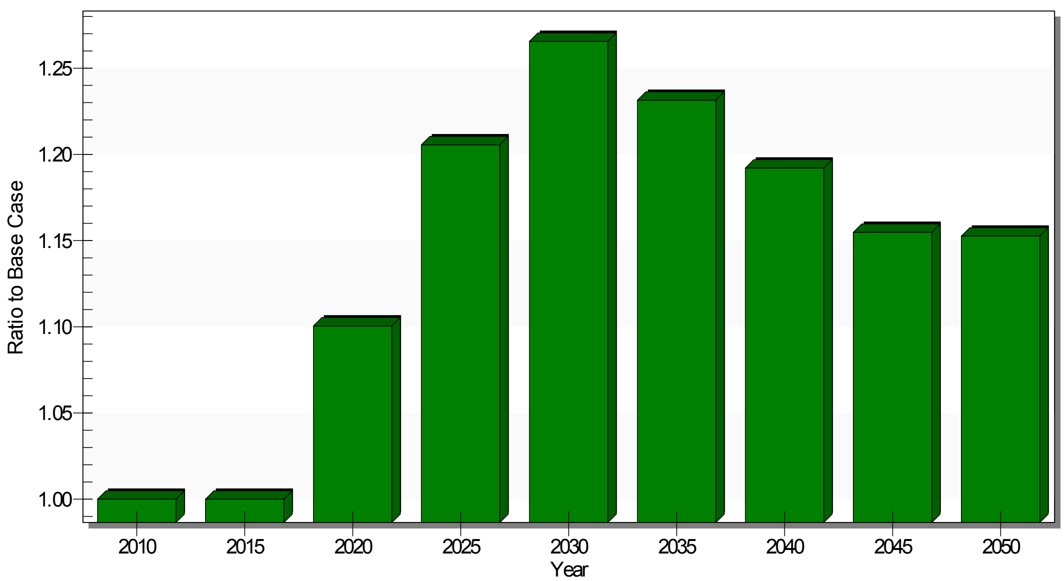

Figure 7.5. Ratio of GDP for Peru in Counterfactual Informal Economy and Base Case scenarios, 2010—2050
Source: IFs version 7.14

[39] All analyses in this section that show results related to the shadow economy required changing the parameter ***gdpshadowon*** from its default value of 0 in the IFs Base Case scenario to 1, so the the model would use the variable GDPSHADOWSHR in forward linkage calculations. Similarly, comparisons of results of such scenarios are with what is effectively an alternative Base Case, namely one in which the parameter's value is set to 1 but no interventions are introduced.

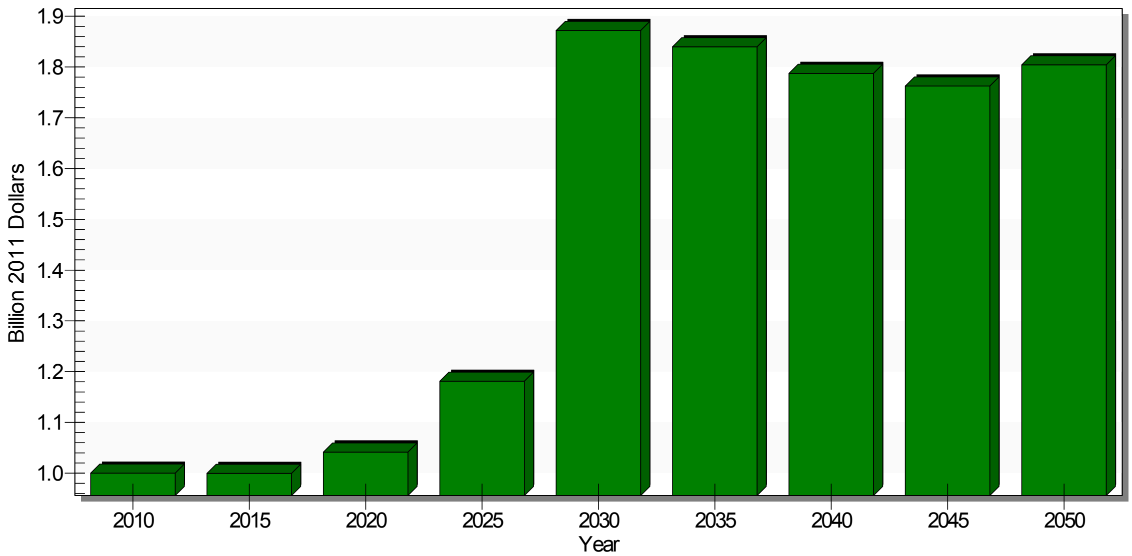

Figure 7.6. Ratio of GDP for Peru in Counterfactual Shadow Economy and Base Case scenarios, 2010—2050
Source: IFs version 7.14

The benefit of the counterfactual in the shadow economy scenario is greater than that in the informal economy scenario. Why? Two factors are at play in each case. Although the per-unit impact of reduction in size of the informal economy is less than the per-unit impact of reduction in size of the shadow economy, the potential magnitude of reduction in the size of the shadow economy is much larger than the potential reduction in the size the informal economy. We have argued in earlier sections that we do not believe the estimates that initialize the sizes of the shadow economy in countries such as Peru. Nonetheless, this analysis helps establish the boundaries of the impact from reduction in informal/shadow economy size and is therefore useful to us.

By the end of the intervention period of the Counterfactual scenario, Peru experiences an 11.5 percentage point decrease in informal GDP relative to the Base Case (roughly 1 standard deviation according to Table 3.3) and a relatively similar or somewhat larger SD reduction in the shadow economy. Over the period from 2016 to 2030 the country averages 1.7 percentage point higher GDP per capita growth rates when compared to the Base Case in Figure 7.5 and 1.2 percentage points higher between 2016 and 2050 when compared to the Base Case in Figure 7.6. These figures are generally similar to those reported by Loayza (2007), who in a study looking at Latin American countries over a 20-year period found empirical evidence indicating that for a one standard deviation increase in the informal economy, countries experienced a 1 to 2 percentage point decrease in per capita GDP growth rates.

However, the Counterfactual scenario obviously only helps us obtain an initial look at the big picture of the informal economy's impact in slowing economic growth (and government revenues). We need to turn to looking at more realistic interventions, singly and in combination, which could, in fact, move workers and economic activity into the formal economy.

Drivers of informality and the impact they might have

As discussed earlier in this report, there is essentially no other analysis of which we are aware of the magnitude of potential impact should a country like Peru be able to reduce informality. We want to undertake that analysis here in two steps. First, we will consider the impact of each driver of informal labor and GDP share individually. Then we will look at the combined impact—and compare that with our counterfactual scenario.

The individual interventions

Earlier sections identified the model's drivers of informal labor and GDP share, how informal labor share is, in turn, linked to informal GDP share, and the forward linkages of those two shares to economic productivity and government revenues. The purpose of the next analysis phase was to learn how individual interventions with respect to the model shares would affect both the informal shares and key forward variables. As Table's 7.1 and 7.2 show, we focus on GDP, government revenues, and poverty levels as the primary impact variables.

It is important to understand the individual interventions made in this sensitivity analysis. The magnitude of each individual intervention was somewhat subjective, but we specified them with an eye to the "aggressive but reasonable" approach of the IFs project and with attention to values in OECD countries around the world. The interventions were:

> *Years of formal adult education.* In Peru, the average adult 15 years of age and older had 8.9 years of formal education in 2010, with males having about 0.6 years more than females. In the Base Case the value for adults is forecast to reach 11.3 years in 2050. In the intervention it reaches 12.5 years, allowing Peru to catch up with Spanish values by 2050, exceeding the value for Germany in 2020 and nearly reaching that of the US in 2010. Because years of adult education is a stock that is slow to change—thanks to adults living many decades beyond their years of formal education—the addition of a year is aggressive. Parameter: ***edyrsagm***
>
> *Government business regulation index*. Although this index officially runs from 1 to 10, country values in 2010 and in our forecasts cluster between 3.5 and 7.5. The value for Peru in 2010 is 5.3 and by 2050 the Base Case forecast for it is 4.7 (lower values are more business-friendly, with Singapore at 3.6 in

2010). In the intervention, this drops aggressively to 3.8 for Peru in 2050. Parameter: ***govbusregindm***

Firm tax rate. Data on this variable are not good. IFs shows the rate to have been 8.5 percent in Peru in 2010, comparable to that in South Korea (8.4 percent), above that of the United Kingdom (8.1 percent), and well above Singapore (7.0 percent). The Base Case forecast for Peru in 2050 is 8.8 percent. The aggressive intervention takes that rate to 6.6 percent in Peru in 2050. Parameter: ***firmtaxrm***

Government to household transfers. In 2010, the Peruvian government spent about 9.6 percent of GDP on combined welfare and pension transfers to households and in the Base Case this rises to 13.2 percent in 2050. Compare this to France which spent 30.7 percent in 2010 or to the United Kingdom at 26.6 percent. In the intervention, the Peruvian values rise an additional 2.2 percentage points to 15.5 percent. Parameters: ***govhhtrnpenm*** and ***govhhtrnwelm***

Government corruption. Using the Transparency International scale from 0 to 10 (prior to their movement to a 100 point scale), higher values are more transparent or less corrupt. In 2010, the value for Peru was 3.5 and in the Base Case rises to 7.9 in 2050, above the level of the United States in 2010. In the very aggressive intervention, the rise is still higher, to 10 at the top of the scale, comparable to Denmark or other Scandinavian countries. Parameter: ***govcorruptm***

Research and development spending. Data for R&D spending for most non-OECD countries is very poor. The IFs value for the portion of GDP spent on R&D in Peru in 2010 is only 0.15 percent (with still less spent, of course, by the government). This contrasts with 2.8 percent in Germany. In the IFs Base Case, Peruvian spending rises to 1.4 percent in 2050, and in the intervention to 2.3 percent. Parameters: **randdexpm** (for public and private spending) and **gdsm**, R&D (for public spending only)

	Labor Informal Share % of Labor			GDP Informal Share % of GDP			GDP Billion $			Government Revenue % of GDP			Poverty Millions		
	'21	**'30**	**'50**	**'21**	**'30**	**'50**	**'21**	**'30**	**'50**	**'21**	**'30**	**'50**	**'21**	**'30**	**'50**
Base Case 2010	69.8			19.0			160.2			20.6			18.8		
Base Case 2021—2050	58.1	48.1	27.3	15.8	12.3	3.4	302	527	1496	20.5	21.2	23.7	15.4	11.3	4.8
Counterfactual GDP share scenario	57.5	44.0	16.3	9.5	1.0	1.0	339	667	1723	21.6	23.7	34.6	14.4	9.1	3.3
Adult Education 1.2 year higher	57.6	46.4	24.1	15.7	11.9	3.0	303	535	1578	20.5	21.3	23.8	15.3	11.1	4.5
Business Regulation 0.9 points lower	55.5	39.9	15.4	15.3	10.6	1.9	305	544	1541	20.6	21.4	23.9	15.3	11.1	4.6
Firm Tax Rate 2.2% lower	57.6	46.4	24.6	15.7	11.9	3.1	303	530	1501	20.3	20.9	23.4	15.4	11.3	4.9
Government Transfers 2.3% higher	57.3	45.7	23.1	15.6	11.8	2.9	303	531	1493	20.6	22.2	25.1	15.3	11.0	4.6
Government Corruption 2.1 points higher	58.1	47.9	26.7	15.1	9.3	3.3	307	568	1662	20.6	21.6	24.0	15.3	10.7	4.0
R&D Spending 0.9% higher	58.1	48.0	27.0	15.6	11.3	3.4	304	541	1577	20.5	21.3	23.8	15.3	11.1	4.5
Combined Intervention All (no counterfactual)	53.6	33.3	11.4	14.0	4.8	1.4	313	629	1889	20.6	22.2	24.9	15.0	9.6	2.8

Table 7.1. Impacts of interventions made for Peru with respect to the informal economy
Note: except in the first line, which shows Base Case values for 2010, values are for 2021, 2030, and 2050 respectively. The poverty rate selected was $10 per day ($2011 PPP)—a level sometimes identified as the entry point to middle-class status for individuals.
Source: IFs version 7.14

	Labor Informal Share % of Labor			GDP Shadow Share % of GDP			GDP Billion $			Government Revenue % of GDP			Poverty Millions		
	‘21	**‘30**	**‘50**	**‘21**	**‘30**	**‘50**	**‘21**	**‘30**	**‘50**	**‘21**	**‘30**	**‘50**	**‘21**	**‘30**	**‘50**
Base Case 2010		69.8			52.8			160.2			20.6			18.8	
Base Case 2021—2050	58.1	48.5	27.3	47.1	41.7	26.7	304	534	1514	20.9	22.1	25.1	15.3	11.2	4.8
Counterfactual GDP share scenario	57.5	44.0	11.4	29.1	1.3	1.0	319	987	2699	23.8	30.3	73.6	14.9	4.6	0.9
Adult Education 1.2 year higher	57.6	46.4	24.4	47.1	41.5	24.4	304	541	1605	20.9	22.1	25.4	15.3	11.0	4.2
Business Regulation 0.9 points lower	55.5	39.9	15.9	46.9	40.3	15.9	305	547	1529	21.0	22.3	25.7	15.3	11.0	4.7
Firm Tax Rate 2.2% lower	57.6	46.4	25.0	47.1	41.5	25.0	304	536	1508	20.7	21.5	25.0	15.3	11.3	4.9
Government Transfers 2.3% higher	57.3	45.7	23.5	47.0	41.4	23.5	304	536	1500	21.0	22.7	27.4	15.3	10.9	4.5
Government Corruption 2.1 points higher	58.1	47.9	26.9	46.3	38.5	22.9	305	547	1709	21.0	22.6	25.0	15.3	11.0	3.8
R&D Spending 0.9% higher	58.1	48.0	27.3	46.9	40.7	23.4	304	540	1604	21.0	22.3	25.2	15.3	11.1	4.3
Combined Intervention All (no counterfactual)	53.6	33.3	11.4	45.6	34.2	11.4	307	573	1926	21.0	23.1	27.1	15.2	10.3	2.7

Table 7.2. Impacts of interventions made for Peru with respect to the shadow economy
Note: except in the first line, which shows Base Case values for 2010, values are for 2021, 2030, and 2050 respectively. The poverty rate selected was $10 per day ($2011 PPP)—a level sometimes identified as the entry point to middle-class status for individuals.
Source: IFs version 7.14

Again, Tables 7.1 and 7.2 shows the individual impacts of each of these interventions. For reducing the size of the informal economy itself (looking at either labor or GDP informal shares), the biggest payoffs are associated with the government to household transfers and business regulation interventions. For reducing the size of the shadow economy, the biggest payoffs are associated with government corruption and business regulation interventions. The biggest payoffs for GDP are, however, corruption reduction, followed by education and R&D spending for the informal economy and shadow economy alike. This begins to point to an important conclusion of this intervention analysis, namely that the interventions work not just through their impact on the size of the informal economy, but also through their additional paths of impact within the larger socioeconomic system. To see this more clearly we can look at the combined interventions.

Combined interventions and conclusion

Figure 7.7 shows three IFs forecasts of GDP per capita in Peru. The Base Case is itself fairly optimistic, with a 3.3 percent compound growth rate. The Combined Intervention scenarios for the shadow economy and informal economy, however, add $10,500 and $9,600 per capita, respectively, another 25 percent by 2050. As the figure indicates, this is clearly superior to the Counterfactual scenario in which the informal economy is effectively eliminated by brute force starting in 2017 (as it is via combined interventions in the policy-focused scenarios).[40] The reason for the superior performance of the combined intervention package has been indicated, but is very important. The interventions do more than accelerate the ongoing reduction in the extent of informality, they boost the growth of Peru in a variety of direct and indirect (and, via positive feedback loops, reinforcing) ways. The package of combined interventions has, in fact, the character of a strategy that would be aimed at helping Peru avoid the so-called Middle-Income Trap and to navigate the middle-income passage. Developing such a strategy and implementing appropriate policies would prove to be a completely complementary effort to the one targeting reduction in informality—the two goals and action packages can fully support one another.

[40] The brute-force Counterfactual Scenario performs best until the early 2030s simply because of the speed at which we imposed it, over 15 years. The individual interventions were also all phased in over 20 years, and there are also lags for many of them between implementation and pay-off.

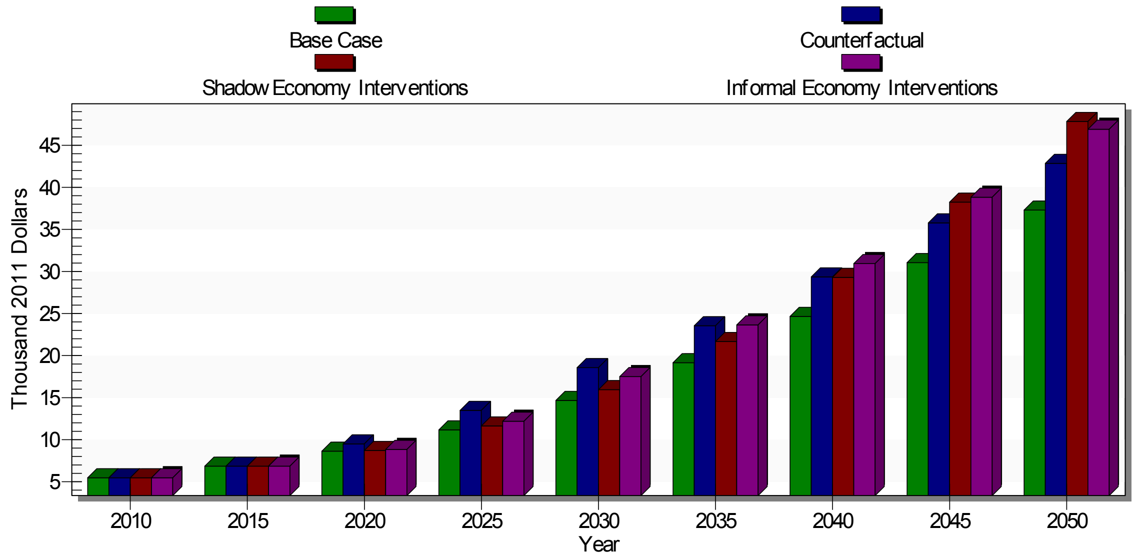

Figure 7.7. GDP per capita (2011$ PPP) in Peru in the Base Case, Counterfactual, and Combined Intervention scenarios
Source: IFs version 7.14

8. Conclusions

This report documents an extensive effort to forecast the size and character of the informal economy of countries globally, with a special focus on Peru. The resulting forecasting capability is, to our knowledge, unique in its ability to explore the long-term future of informality. The study produces forecasts for the informal labor share and the informal GDP share, and augments that that with forecasts for the shadow economy.

This project has studied the conceptualization and data on informal economies, recognizing that there are great variations in conceptualizations and measurement. It has used existing literature and the project's own very extensive empirical analysis to tease out understandings of both the drivers, or determinants, of the informal economy and the implications of informality for economic growth and government finance.

The resulting extension of the International Futures (IFs) modeling system allows the presentation of Base Case forecasts of informal labor and informal and shadow GDP shares—the dynamic path that informality appears to be following. It also facilitates the development of alternative scenarios and therefore the exploration of interventions that might change the future pattern of informality. Countries may seek to do so both because they wish to reduce the social costs of informality and because they understand doing so to be an important part of navigating the middle-income passage to high-income status.

In the specific case of Peru, our own early analysis suggests first that informality is poised to decrease substantially, even in the Base Case scenario. Second, it indicates that an "aggressive but reasonable" intervention package can, in fact, contribute to the interactive accomplishment of both informality reduction and passage to high-income. With such a concerted effort Peru might be able to increase its GDP per capita of 2050 by more than 25 per cent relative to our forecast without it.

All studies, of course, should express caveats clearly. In this case, we must recognize many. They include the uncertainty of basic data (even of conceptualization) and, even more, the complexity of understanding and representing relationships. All models are simplifications of reality and therefore all necessarily not only suffer from potential weaknesses in the relationships included, but from omitting other dynamics. Nonetheless, we stress again both the uniqueness of this project and the extent of the progress made within it. May it help Peru and other countries chart paths to more desirable futures.

Bibliography

Works Cited

Agénor, Pierre-Richard, and Otaviano Canuto. 2012. “Middle-Income Growth Traps.” Policy Research Working Paper no. 6210. World Bank, Washington DC.

Ahmed, Ehsan, J. Barkley Rosser, Jr., and Marina V. Rosser. 2005. “A Global Perspective on the Non-Observed Economy, Inequality, Corruption, and Social Capital.” Department of Economics, James Madison University: 1—61.

Andrei, Tudorel. 2010. “Quantitative Methods for Evaluating the Informal Economy. Case Study at the Level of Romania.” *Theoretical and Applied Economics* XVII (7): 15—24.

Andrews, Dan, Aida Caldera Sánchez, and Asa Johansson. 2011. “Towards a Better Understanding of the Informal Economy.” Economics Department Working Papers no. 873. OECD, Paris.

Aruoba, S Boragan. 2010. “Informal Sector, Government Policy and Institutions.” Society for Economic Dynamics Meeting Papers no. 324. Society for Economic Dynamics, St. Louis MO, USA.

Bacchetta, Marc, Ekkehard Ernst, and Juana P. Bustamante. 2009. *Globalization and Informal Jobs in Developing Countries: A Joint Study of the International Labour Office and the Secretariat of the World Trade Organization*. Geneva: World Trade Organization Publications and International Labour Office Publications.

Bailey, Martin, Diana Ferrell, and Jaana Remes. 2005. “Domestic Services: The Hidden Key to Growth.” McKinsey Global Institute Report. McKinsey & Company.

Barro, Robert, and Jong-Wha Lee. 2010. “A New Data Set of Educational Attainment in the World, 1950—2010.” *Journal of Development Economics* 104: 184—198.

Baser, Didem Dincer, Diana Farrell, and David E. Meen. 2006. “The Cost of the Gray Market in Turkey.” In Diana Farrell ed., *Driving Growth: Breaking Down Barriers to Global Prosperity*. McKinsey Global Institute. Cambridge, MA USA: Harvard Business Press.

Bosch, Mariano, Marco Stampini, and Guadalupe Bedoya. 2012. “Are Conditional Cash Transfers Associated with Lower Labor Supply? Evidence from Mexico’s Oportunidades over the period 2000—2010.” Technical Note no. TN-497. Inter-American Development Bank, Washington DC.

Breusch, Trevor. 2005. "'Estimating the Underground Economy Using MIMIC Models.'" Working Paper, National University of Australia, Canberra.

Buehn, Andreas, and Friedrich Schneider. 2012. "Shadow Economies around the World: Novel Insights, Accepted Knowledge, and New Estimates." *International Tax and Public Finance* 19(1): 139—171. doi:10.1007/s10797-011-9187-7.

———. 2013. "Estimating the Size of the Shadow Economy: Methods, Problems and Open Questions." Working Paper no. 1322. Department of Economics, Johannes Kepler University of Linz, Auhof, Austria.

Centeno, Miguel Angel, and Alejandro Portes. 2006. "The Informal Economy in the Shadow of the State." In Patricia Fernández-Kelly and Jon Shefner eds., *Out of the Shadows: Political Action and the Informal Economy in Latin America*. Pennsylvania: Penn State University Press.

Chen, Martha Alter. 2007. "Rethinking the Informal Economy: Linkages with the Formal Economy and the Formal Regulatory Environment." DESA Working Paper no. 46. United Nations Department of Economics and Social Affairs, New York.

———. 2012. "The Informal Economy: Definitions, Theories and Policies." WIEGO Working Paper no. 1. Women in Informal Employment Globalizing and Organizing, Cambridge, MA.

Chong, Alberto, and Mark Gradstein. 2004. "Inequality, Institutions , and Informality." Research Department Working Paper no. 516. Inter-American Development Bank, New York.

de Brauw, Alan, Daniel O. Gilligan, John Hoddinott, and Shalini Roy. 2013. "Bolsa Família and Household Labour Supply." One Pager no. 239. International Policy Centre for Inclusive Growth, Brasilia, Brazil.

de Soto, Hernado. 1989. *The Other Path: the Economic Answer to Terrorism*. New York: Basic Books.

———. 2000. *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. New York: Basic Books.

Dell'Anno, Roberto, and Friedrich Schneider. 2006. "Estimating the Underground Economy by Using MIMIC Models: A Response to T. Breusch's Critique." Working Paper no. 0607. Department of Economics, Johannes Kepler university of Linz, Auhof, Austria.

Dickson, Janet R, Barry B Hughes, and Mohammod T Irfan. 2010. *Advancing Global Education*. Vol. 2 of the Patterns of Potential Human Progress series. Boulder, CO, and New Delhi, India: Paradigm Publishers and Oxford University Press.

Docquier, Frédéric, Tobias Müller, and Joaquín Naval. 2014. “Informality and Long-Run Growth.” IZA Discussion Paper no. 7883. Institute for the Study of Labor, Bonn, Germany.

Ela, Mehmet. 2013. “An Assessment on the Relationship Between Informal Economy and Educational Level in Turkey.” *International Journal of Economics and Financial Issues* 3(4): 910—922.

Elgin, Ceyhun, and O. Oztunali. 2012. “Shadow Economies around the World: Model Based Estimates.” Working Papers no. 2012/05. Department of Economics, Bogazici University, Istanbul.

Elgin, Ceyhun, and Friedrich Schneider. 2013. “Shadow Economies in OECD Countries: DGE vs. MIMIC Approaches.” Working Papers no. 2013/13. Department of Economics, Bogazici University, Istanbul.

Enste, Dominik. 2009. “Regulation and shadow economy: empirical evidence for 25 OECD-countries.” *Constitutional Political Economy* 21: 231—248. doi: 10.1007/s10602-009-9081-9

Eurostat. 2013. *Essential SNA: Building the Basics*. 2013 Edition. Methodologies & Working Papers. Luxembourg: Publications Office of the European Union.

Feige, Edgar L. 1990. “Defining and Estimating Underground and Informal Economies: The New Institutional Economic Approach.” *World Development* 18(7): 989—1002

Friedman, Eric, Simon Johnson, Daniel Kaufmann, and Pablo Zoido-Lobaton. 2000. “Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries.” *Journal of Public Economics* 76: 459—493. doi:10.1016/S0047-2727(99)00093-6.

Furnival, George M. and Robert W. Wilson. 1974. “Regressions by Leaps and Bounds.” *Technometrics* 16(4): 499—511

García-Verdú, Rodrigo. 2007. “Measurement of the Shadow Economy or Shadowy Measurement?” Draft. 1–36.

Gërxhani, Klarita. 2004. “The Informal Sector in Developed and Less Developed Countries: A Literature Survey.” *Public Choice* 120(3/4): 267—300. doi:10.1023/B:PUCH.0000044287.88147.5e.

Gyomai, György. 2012. "Summary of the OECD Survey on Measuring the Non-Observed Economy." Working Party on National Accounts. OECD, Paris.

Gyomai, György, and Peter van de Ven. 2014. "The Non-Observed Economy in the System of National Accounts." Working Party on National Accounts. OECD, Paris.

Hart, Keith. 1973. "Informal Income Opportunities and Urban Employment in Ghana." *Journal of Modern African Studies* 11(1): 61—89. DOI: http://dx.doi.org/10.1017/S0022278X00008089

Heintz, James, and Robert Pollin. 2003. "Informalization, Economic Growth and the Challenge of Creating Viable Labor Standards in Developing Countries." Working Paper Series no. 60. Political Economy Research Institute, University of Massachusetts, Amherst, Amherst MA.

Instituto Nacional de Estadistica e Informatica (INEI). 2014. "Satellite Account of the Informal Economy: Production and Informal Employment in Peru 2007—2012." Working Document 7/3/2014. INEI, Lima, Perú.

International Labor Organization (ILO). 2002. *Women and Men in the Informal Economy: A Statistical Picture.* First Edition. Geneva: International Labour Office.

———. 2012. "Statistical Update on Employment in the Informal Economy." Statistical Update. ILO Department of Statistics, International Labour Office, Geneva.

———. 2013a. *Measuring Informality: A Statistical Manual on the Informal Sector and Informal Employment*. Geneva, Switzerland: International Labour Office.

———. 2013b. *Women and Men in the Informal Economy: A Statistical Picture.* Second Edition. Geneva: International Labour Office.

———. 2014a. "New Enhanced Database on Informal Employment Produced by the ILO in Partnership with WIEGO." http://www.ilo.org/global/statistics-and-databases/WCMS_179795/lang--en/index.htm.

———. 2014b. Panorama Temático Laboral: Transición a la Formalidad en América Latina y el Caribe. Lima, Peru: Organización International del Trabajo.

Jütting, Johannes, and Juan R. de Laiglesia. 2009. *Is Informal Normal? Towards More and Better Jobs in Developing Countries*. Paris: OECD Development Centre. doi:10.1787/9789264059245-en.

Kanchoochat, Veerayooth, and Patarapong Intarakumnerd. 2014. "Tigers Trapped: tracing the Middle-Income Trap through the East and Southeast Asian Experience. Berlin Working Papers on Money, Finance, Trade, and Development no. 4. Competence Center, Berlin.

Kim, Byung-yeon. 2005. "Poverty and Informal Economy Evidence from Romania." *Economics of Transition* 13(1): 163—185.

Lawless, J. F. and K. Singhal. 1978. "Efficient Screening of Nonnormal Regression Models." *Biometrics* 34(2): 318—327

Lindsey, Charles and Simon Sheather. 2014. "Stata module to perform best subsets variable selections." Statistical Software Components no. S457816. Department of Economics, Boston College.

Lisi, Gaetano, and Maurizio Pugno. 2011. "The Underground Economy in a Matching Model of Endogenous Growth." MPRA Paper no. 31312.

Loayza, Norman V, Ana María Oviedo, and Luis Servén. 2005. "The Impact of Regulation on Growth and Informality: Cross-Country Evidence." Working Paper Series no. 3623. World Bank, Washington DC.

———. 2007. "The Causes and Consequences of Informality in Peru." Working Paper Series 2007-018. Banco Central De Reserva Del Perú, Lima, Peru.

Macias, Jose Brambila. 2008. "Modeling the Informal Economy in Mexico: A Structural Equation Approach." MPRA Paper no 8504. Munich Personal RePEc Archive, Munich University, Germany.

OECD. 2002. *Measuring the Non-Observed Economy: A Handbook*. OECD Publishing. doi:10.1787/9789264175358-en.

Oviedo, Ana Maria, Mark R. Thomas, and Kamer Karakurum-Özdemir. 2009. "Economic Informality: Causes, Costs, and Policies: A Literature Survey." World Bank Working Paper no. 167. World Bank, Washington DC.

Perry, Guillermo E, William F Maloney, Omar S Arias, Pablo Fajnzylber, and Andrew D Mason Jaime Saavedra-Chanduvi. 2007. *Exit and Exclusion*. Vol. 57. Washington DC, USA: The World Bank.

Schneider, Friedrich. 2005. "Shadow Economies around the World: What Do We Really Know?" *European Journal of Political Economy* 21(3): 598—642. doi:10.1016/j.ejpoleco.2004.10.002.

———. 2011. "The Shadow Economy Labour Force What Do We (not) Know?" *World Economics* 12(4): 1—40.

———. 2012. “The Shadow Economy and Work in the Shadow: What Do We (Not) Know?” IZA Discussion Paper no. 6423. Institute for the Study of Labor, Bonn, Germany.

———. 2013. *The Shadow Economy*. London: The Institute of Economic Affairs. doi:10.1017/CBO9781139542289.

Schneider, Friedrich, Andreas Buehn, and Claudio E. Montenegro. 2010. “New Estimates for the Shadow Economies All over the World.” *International Economic Journal* 24(4): 443—461. doi:10.1080/10168737.2010.525974.

Schneider, Friedrich, and Dominik H Enste. 2000. “Shadow Economies: Size, Causes, and Consequences.” *Journal of Economic Literature* 38: 77—114. doi:10.1257/jel.38.1.77.

Schneider, Friedrich, and Colin Williams. 2013. *The Shadow Economy*. London: The Institute of Economic Affairs. doi:10.1017/CBO9781139542289.

Tahir, Nadia. 2014. “Informal Employment in Pakistan: Survivalist or Structuralist?” *ATM-2013 Conference*, ERUDITE University of Paris est Marne-la-Vallée.

Tansel, Aysit, and Elif Oznur Kan. 2012. “Labor Mobility across the Formal/Informal Divide in Turkey: Evidence from Individual Level Data.” IZA Discussion Paper Series no. 6271. Institute for the Study of Labor, Bonn, Germany.

Tanzi, Vito. 1999. “Uses and Abuses of Estimates of the Underground Economy.” *The Economic Journal* 109: F338—F347. doi:10.2307/2566007.

Tho, Tran Van. 2013. “The Middle-Income Trap: Issues for Members of the Association of Southeast Asian Nations.” ADBI Working Paper Series no 421. Asian Development Bank Institute, Tokyo.

Torgler, B., and Friedrich Schneider. 2007. “Shadow Economy, Tax Morale, Governance and Institutional Quality: A Panel Analysis.” IZA Discussion Paper Series no. 2563. Institute for the Study of Labor, Bonn, Germany.

Torrini, Roberto. 2005. “Cross-Country Differences in Self-Employment Rates: The Role of Institutions.” *Labour Economics* 12: 661—683. doi:10.1016/j.labeco.2004.02.010.

Statistical Consulting Group. No date. Stata Web Books: Regression with Stata. Statistical Consulting Group, University of California at Los Angeles. Online at: http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm [accessed 6/8/15]

United Nations Economic Commission for Europe (UNECE). 2008. *Non-Observed Economy in National Accounts: Survey of Country Practices*. New York and Geneva: UNECE.

Vuletin, Guillermo. 2008. "Measuring the Informal Economy in Latin America and the Caribbean." IMF Working Paper no. 08/102. International Monetary Fund, Washington DC.

World Bank. 2014a. "Informal Employment." *Gender Statistics Database*. Available at http://databank.worldbank.org/. [Accessed on 3/35/2015]

———. 2014b. *Doing Business 2015: Going Beyond Efficiency.* Washington, DC: World Bank. License: Creative Commons Attribution CC BY 3.0 IGO DOI: 10.1596/978-1-4648-0351-2

———. 2008. Peru: Trajectories Towards Formality. Report no. 43737-PE. Bolivia, Ecuador, Peru and Venezuela Country Management Unit. Lima, Peru: World Bank.

Additional Sources Consulted

Albu, Lucian-Liviu. 2007. "A Model to Estimate Informal Economy at Regional Level: Theoretical and Empirical Investigation." Munich Personal RePEc Archive Paper no. 3760. Institute for Economic Forecasting, Bucharest.

Albu, Lucian-Liviu, Raluca Lorgulescu, and Cristian Stanica. 2010. "Estimating Hidden Economy and Hidden Migration: The Case of Romania." *Romanian Journal of Economic Forecasting* 2(13): 46—56.

Andrei, Tudorel, Andreea Iluzia Iacob, Stelian Stancu, and Bogdan Oancea. 2010. "Quantitative Techniques Used for the Informal Economy Analysis at National and Regional Levels." *Informatica Economica* 14(3): 153—165.

Arvin-rad, Hassan, Arnab K. Basu, and Maria Willumsen. 2010. "Economic Reform, Informal-Formal Sector Linkages and Intervention in the Informal Sector in Developing Countries: A Paradox." Discussion Paper no. 5229.

Avirgan, Tony, L. Josh Bivens, and Sarah Gammage, eds. 2005. *Good Jobs, Bad Jobs, No Jobs: Labor Markets and Informal Work in Egypt, El Salvador, India, Russia, and South Africa*. Washington DC: Economic Policy Institute.

Bivona, Enzo, Francesco Ceresia, and Giovan Battista Montemaggiore. 2006. "Exploring the Underground Economy through System Dynamics to Support Public Decision Makers: a preliminary qualitative analysis." 1—17.

Bosch, Mariano, and William F. Maloney. 2010. “Comparative Analysis of Labor Market Dynamics Using Markov Processes: An Application to Informality.” *Labour Economics* 17(4): 621—631. doi:10.1016/j.labeco.2010.01.005.

Buehn, Andreas, and Friedrich Schneider. 2013. “Shadow Economies in Highly Developed OECD Countries: What Are the Driving Forces?” IZA Discussion Papers no. 6891. Institute for the Study of Labor, Bonn, Germany.

Burov, Vitaliy Yurievich. 2012. “Theoretical Aspects of Shadow Economy Study.” *Advanced Research in Scientific Areas* 2012: 263—265.

Chambwera, Muyeye, James Macgregor, and Antonia Baker. 2011. “The Informal Economy: A Primer for Development Professionals on the Importance of the Informal Economy in Developing Countries.” London: International Institute for Environment and Development.

Davies, Rob, and James Thurlow. 2010. “Formal-Informal Economy Linkages and Unemployment in South Africa.” *South African Journal of Economics* 78(4): 437—459. doi:10.1111/j.1813-6982.2010.01250.x.

Elgin, Ceyhun. 2010. “Political Turnover, Taxes and the Shadow Economy.” Working Papers no. 2010/08. Department of Economics, Bogazici University, Istanbul.

Fernández-Kelly, Patricia, and Jon Shefner. 2006. *Out of the Shadows: Political Action and the Informal Economy in Latin America*. University Park, Pennsylvania: Pennsylvania State University Press.

Georgiou, George M. 2007. “Measuring the Size of the Informal Economy: A Critical Review.” Working Paper Series 2007-1. Central Bank of Cyprus, Nicosia, Cyprus.

Gërxhani, Klarita, and Herman G. van de Werfhorst. 2013. “The Effect of Education on Informal Sector Participation in a Post-Communist Country.” *European Sociological Review* 29(3): 464—476. doi:10.1093/esr/jcr087.

Godfrey, Paul C. 2011. “Toward a Theory of the Informal Economy.” *The Academy of Management Annals* 5(1): 231—277. doi:10.1080/19416520.2011.585818.

Inter-American Development Bank. 2003. *Good Jobs Wanted: Labor Markets in Latin America*. Economic and Social Progress Report 2004. Washington DC: Inter-American Development Bank.

Johnson, Simon, Daniel Kaufmann, and Andrei Shliefer. 1997. “The Unofficial Economy in Transition.” *Brookings Papers on Economic Activity* 2: 159—239. doi:10.1093/oxfordhb/9780199759927.013.0032.

Kus, Basak. 2006. "State and Economic Informality in a Comparative Perspective." CCOP Working Papers no. 2006-11. Center for Culture, Organization, and Politics, University of California, Berkley.

Lippert, Owen, and Michael Walker. 1997. *The Underground Economy: Global Evidence of Its Size and Impact*. Vancouver, British Colombia, Canada: The Fraser Institute.

Loayza, Norman V. 1996. "The Economics of the Informal Sector: A Simple Model and Some Empirical Evidence from Latin America." *Carnegie-Rochester Conference Series on Public Policy* 45: 129—162. doi:10.1016/S0167-2231(96)00021-8.

Farrell, Diana, ed. 2006. *Driving Growth: Breaking Down Barriers to Global Prosperity*. McKinsey Global Institute. Cambridge, MA USA: Harvard Business Press.

Meagher, Kate. 2013. "Unlocking the Informal Economy: A Literature Review on Linkages Between Formal and Informal Economies in Developing Countries." WIEGO Working Paper no. 27. Women in Informal Employment Globalizing and Organizing, Cambridge, MA USA.

Olaya, Camilo, Fabio Díaz, and Santiago Caicedo. 2007. "Towards a System Dynamics Model of De Soto's Theory on Informal Economy." *Proceedings of the 25th International Conference of the Systems Dynamics Society:* 1-20.

Porta, Rafael La, and Andrei Shleifer. 2008. "The Unofficial Economy and Economic Development." *Brookings Papers on Economic Activity* 2008: 275—363. doi:10.1353/eca.0.0016.

Portes, Alejandro and William Haller. 2005. "The Informal Economy." In Neil j. Smelser & Richard Swedberg eds., *The Handbook of Economic Sociology*. Princeton, NJ: Princeton University Press.

Safa, Helen I. 1986. "Urbanization, The Informal Economy And State Policy In Latin America." *Urban Anthropology and Studies of Cultural Systems and World Economic Development* 15(1): 135—163.

Schramm, Heidi R. 2014. "The Equilibrium Effects of Income Taxation on Formal and Informal Labor Markets." NEUDC 2014 Working Paper. Northeast Universities Development Consortium Conference, Boston University.

Teixeira, Clarissa Gondim. 2009. "What Is the Impact of Cash Transfers on Labour Supply?" One Pager no. 85. International Policy Centre for Inclusive Growth, Brasilia, Brazil.

Appendix A1: Summary of Informal Economy Conceptualizations

An expanded look at the different conceptualizations of the informal economy developed in the literature.

Table A1.1: Full summary of informal economy conceptualizations						
Source	Label/Focus	Definition	Components	Estimation Methods	Measures/Data	Criticisms
Gyomai and van de Ven 2014; Eurostat 2013; UNECE 2008	Non-observed economy Focus: adjusting formal GDP to capture informal GDP	Refers to all productive activities that may not be captured in the basic data sources used for compiling national accounts whether deliberately or through data deficiencies. An umbrella term that encompasses five major components.	**Underground production** (legal activities deliberately concealed); **illegal production**; **informal sector production** (activities by unincorporated/unregistered enterprises below a certain level of employment); **household production for own final use** (goods/services produced by the households that consume them); and **non-observed informal activities** (activities that should be accounted for but are missed due to deficiencies in the statistical programs)	Direct statistical methods (surveys on expenditure, income, labor, etc.), indirect statistical methods (supply-based approach, demand-based approach, income-based approach, commodity flow approach).	Agricultural, business, household, demographic surveys and censuses, taxation and fiscal data, household budget and expenditure data.	Non-timely nature and sparse coverage from being survey based, methodology and data differ across countries, the 5 categories have blurred edges.
Elgin and Schneider 2013; Schneider 2012	Shadow economy Focus: calculating size of	All market-based legal production of goods and services that would contribute to officially calculated	Includes: **Underground production; informal production** Excludes: **illegal activities**, **household own final use**	MIMIC model and Physical input (electricity consumption) method, and currency demand method	Personal income taxes, indirect taxation, tax morale, social security contributions, regulation, quality	MIMIC only produces relative estimates of size and development of the shadow economy, currency demand needed to

	informal GDP	gross national product but are deliberately concealed from public authorities.			of state institutions, unemployment, self-employment, and GPD per capita	convert to absolute values as % of GDP,
ILO 2013a; Eurostat 2013; Chen 2012; World Bank 2008	Informal sector (enterprises only) Focus: informal enterprises' contribution to formal GDP	All activities by unincorporated household enterprises that do not conform to the exiting legal and administrative framework, particularly the lack of a pension (legal definition of informality) and are below a certain number of employees with low capital and low productivity (production definition of informality).	Includes: **underground production, illegal production** (self-employed, family units, and micro-entrepreneurs) Excludes: **household own final use, agricultural activities**	Direct methods (surveys on expenditure, firm income), indirect statistical methods (supply-based approach, demand-based approach, income-based approach, commodity flow approach)	Business surveys, taxation and fiscal data	The small size requirement for the production definition is not sufficient unless used in combination with other criteria since size limit is more or less arbitrary. The legal definition is based on official registration data which may pose problems for comparability between countries, and the practical problems of keeping up to date registries.
ILO 2013a; ILO 2013b; Eurostat 2013;	Informal economy (labor only) Focus: informal	Comprises informal employment (without secure contracts, worker benefits, or social protection) both	**Informal employment in informal enterprises** (employers, employees, own-account operators, and unpaid family workers in informal enterprises)	Uses residual method--indirect approach to estimating total informal employment and major components	total non-agricultural workforce; number of formal employees; calculated from data: total informal	Persons engaged in very small-scale or casual self-employment activities may not report in statistical surveys that they

Chen 2012	labor as a share of total labor	inside and outside informal enterprises.	**Informal employment outside informal enterprises** (domestic workers, casual or day laborers, temporary or part-time workers industrial out-workers and unregistered or undeclared workers)	outside of agriculture using data from population census, labor force surveys, etc.	employment; paid employees and self-employed; informal paid employment; total informal employment inside and outside informal enterprises	are self-employed or employed at all; informal sector statistics may be affected by efforts in classifying certain groups whose activity is on the borderline between self-employment and wage employment, the enterprise based definition of the informal sector is unable to capture all aspects of the increasing informalization of employment.
Andrei et al. 2010	Hidden economy Focus: calculating size of informal GDP	Economic activities and transactions that deliberately avoid existing legal frameworks	Economic actors who elude the legal framework to avoid paying taxes or duties or to avoid certain bureaucratic barriers	Uses econometric approach to assess the proportion of total cash outside the banking system used in the hidden economy	Cash outside the banking system, total government expenses, GDP, short-term interest rate, inflation rate, taxes on products	

Appendix A2: Summary of Methods for Estimating Informality

An expanded look at the different methods for estimating informality used in the literature.

Table A2.1 Methods for Estimating Informality				
Method	Description	Strengths	Weaknesses	Sources
		Direct Methods		
Direct surveys, tax audits	Microeconomic approaches using surveys and samples based on voluntary replies, tax auditing, and other compliance methods to measure the informal economy.	Can provide great level of detail about the structure of the informal economy.	Precision and results depend on questionnaire formulation and respondents' willingness to cooperate. Tax audits do not provide a random sample.	Buehn and Schneider 2012; Macias 2008; Vuletin 2008
		Indirect Methods		
Discrepancy between national expenditure and income	Differences between national income and national expenditure estimates are used to approximate the size of the informal economy.	Would be a particularly strong measure if the components of expenditure could be measured without error or omission.	Discrepancy can reflect other types of omissions and errors and some estimates of expenditure are not statistically independent from income calculations.	Buehn and Schneider 2012; Vuletin 2008
Discrepancy between official and actual labor force	A decline in official labor force participation is seen as an increase in the size of the informal economy, assuming total participation is constant.		Changes in the participation rate may have many other explanations: position in business cycle, difficulty finding a job, decisions around education and retirement.	Vuletin 2008
Currency Demand	Assuming informal transactions are cash-based, a change in the size	One of the most commonly used approaches. Also allows	Assumes a constant demand for currency across countries.	Buehn and Schneider 2012; Andrews et

	of the informal economy will impact the demand for currency. Uses a money demand equation to estimate demand over time.	comparison of size of informal economy to official GDP.	Assumes transactions in the informal economy only occur in cash. Most studies assume tax burden only cause.	al. 2011; Vuletin 2008
Electricity Consumption	Uses electricity consumption as proxy for total economic activity. Size of informal economy based on the difference in growth in electricity consumption compared with the difference in GDP growth.	Very simple method and uses very well established relationship between GDP and electricity consumption.	Not all informal activities use electricity. Elasticity of electricity use to economic growth can vary significantly between countries and even within countries. Technology advance increases the efficiency of electricity use for both formal and informal activities.	Buehn and Schneider 2012; Andrews et al. 2011; Vuletin 2008
Transaction Approach	Based on assumption of constant relation over time between the volume of transactions and official GNP. Size of informal economy calculated by subtracting official GNP from total nominal GNP.		Requires a base year where informal economy size is 0. Large amount of data required to eliminate financial cross payments, velocity of cash transactions hard to track.	Buehn and Schneider 2012
Model Approaches				
Multiple indicators, multiple causes	Simultaneous model that uses relationships between observable	Uses multiple indicators, can provide cross-country,	Results are highly sensitive to changes in units and input variable	Buehn and Schneider 2013; Andrews et

estimation (MIMIC)	variables (causes and effects) to estimate value of the latent, unobserved value (the informal economy). A factor analytic approach is used to measure the size of the informal economy over time.	multiyear analysis.	transformations, theoretical foundation for variables to include/exclude often weak, definition of latent variable often subjective.	al. 2011; Macias 2008; Breusch 2005
Dynamic general equilibrium model (DGE)	Two-sector dynamic equilibrium model using national income statistics to calculate size the informal economy in % of GDP.	Uses multiple indicators, can provide cross-country, multiyear analysis.	Reliance on national income statistics limits the number of variables that might affect informal economy size.	Elgin and Schneider 2013

Appendix A3: Definitions of Informal Sector Enterprises and Labor used by selected countries

A sample of the definitions used by individual countries in compiling statistics under the non-observed economy framework developed by the OECD.

Definitions of informal sector enterprises and labor used by selected countries		
Country	Source	Definition
Brazil	Urban informal economy survey, national household sample survey	**Enterprises:** household unincorporated enterprises with fewer than six employees and without a complete set of accounts (agriculture excluded). **Labor:** employees without a formal contract
Mexico	National survey of occupations and employment	**Enterprises:** household unincorporated enterprises that have no complete set of accounts and are not registered (agriculture excluded). **Labor:** employees without access to public or private health services by virtue of their job
Panama	Household survey	**Enterprises:** household unincorporated enterprises with fewer than five employees (agriculture excluded). **Labor:** employees without employment contract, plus employees with employment contract who are not covered by social security as directly insured persons (excluding retirees)
Mali	Labor force survey, 2007	**Enterprises:** private enterprises with fewer than 11 persons engaged that are not registered with the National Institute for Social Protection and do not have accounts (agriculture excluded). **Labor:** employees for whom the employer does not pay social contributions and who are not entitled to paid annual and sick leave.
Republic of Moldova	Labor force survey	**Enterprises:** household unincorporated enterprises that are not registered (agriculture included). **Labor:** employees for whom the employer does not pay social contributions, or who do not benefit from paid annual leave (or financial compensation for untaken leave), or who will not be given paid sick leave in the event of illness or injury
Russian Federation	Population survey on employment problems	**Enterprises:** household unincorporated enterprises that are not registered as a legal entity or have no legal status (agriculture included). **Labor:** employees without a labor

		contract
Turkey	Household labor force survey	**Enterprises:** household unincorporated enterprises paying a lump-sum tax or not paying any tax, and with fewer than 10 persons engaged (agriculture excluded). **Labor:** employees without any social security registration
India	National sample survey, 61st Round (2004-2005)	**Enterprises:** household unincorporated enterprises with fewer than 10 persons engaged (agriculture excluded). **Labor:** employees not entitled to social security benefits or paid sick or annual leave (agriculture excluded)
Source: ILO 2013: 23; 41		

Appendix A4: Statistical Justification for Model Specification

This section expands upon the discussion in section 4 of the report by describing the statistical methods used to specify and justify the models representing the informal economy in the IFs system. The task of developing the greater representation was not an exercise of pure econometric analysis; instead it also draws heavily from a system-level understanding of informality – namely that informal production is dependent on an informal labor presence. Due to its important role as a driving variable of informal GDP, and because of its relative strong empirical base, we prioritize the specification of the informal labor model and then consider the output as a potential driving variable for the subsequent two models. Furthermore, due to the endogenous interdependence of forecast variables within the IFs model we prohibit multiple occurrence of driving variables across the models of informal labor and GDP.[41]

Informal labor as a percent of total labor

Starting with the full set of variables listed at the beginning of Section 4 we began to explore best sub-set specification options. This method applies the Furnival-Wilson (1975) leaps-and-bounds algorithm, "using the log likelihoods of candidate models, allowing variable selection to be performed on a wide family of normal and non-normal regression models" (Lindsey and Sheather 2014). Further details regarding this method are discussed in (Lindsey and Sheather 2014) and Lawless and Singhal (1978). Table A.1 reports the log likelihood (LL), Akaike's information criterion (AIC), and the Bayesian information criterion (BIC) to assess the best regressions for each predictor sub-set.[42]

[41] Note: In Section 4 we reported the regression outputs using all available observations. For this reason, the coefficients found in the final models of Section 4 are those that have been used to operationalize our models of informality within the IFs system. In this appendix we have limited the observations to those that are shared by all possible explanatory variables in order to ensure comparability between model options using a variety of validation techniques. Therefore, the coefficients reported in this section may differ slightly from those found in Section 4 and in the IFs model.

[42] The rows in Table A.1 correspond to the following variable subsets: (1) Gov Trans to HH; (2) Business Reg, Gov Trans to HH; (3) Education, Business Reg, Gov Trans to HH; (4) Education, Business Reg, Tax Rate, Gov Trans to HH; (5) Education, Business Reg, Tax Rate, Gov Trans to HH, Gini; (6) Education, Business Reg, Tax Rate, Spend on R&D, Gov Trans to HH, Gini; (7) Education, Business Reg, Tax Rate, Corruption, Spend on R&D, Gov Trans to HH, Gini; (8) Education, GDP per capita, Business Reg, Tax Rate, Corruption, Spend on R&D, Gov Trans to HH, Gini; (9) Education, GDP per capita, Business Reg, Tax Rate, Corruption, Spend on R&D, Gov Trans to HH, Gini, Labor Growth; (10) Education, GDP per capita, Business Reg, Tax Rate, Corruption, Spend on R&D, Gov Trans to HH, Gov Effect, Gini, Labor Growth

# of IVs	LL	AIC	BIC
1	-139.94	283.88	286.93
2	-130.28	266.56	271.14
3	-127.53	263.07	269.18
4	-125.23	260.46	268.09
5	-123.42	258.84	**268.00**
6	-121.72	**257.44**	268.13
7	-121.14	258.27	270.49
8	-121.00	260.00	273.74
9	-120.88	261.77	277.03
10	-120.87	263.74	280.53

Table A.1: Best subset results for the informal labor model. The best models identified by the AIC and BIC values have been bolded. Models specific to each of these sub-sets are listed in Footnote 42.
Source: IFs version 7.14

Initial results indicated consistent significance and stability in the education, regulation, taxation, and government transfers variables. By comparing the relative quality of the models using Akaike information criterion (AIC) and the Bayesian information criterion (BIC) two models were identified as best subsets from the universe of explanatory variables.

	Model 1	Model 2	Model 3
Education	-3.514**	-3.080**	-2.831*
	(-3.32)	(-2.91)	(-2.60)
Business Reg	10.86**	12.49**	12.55**
	(3.02)	(3.50)	(3.39)
Tax Rate	10.89*	9.392*	9.185*
	(2.55)	(2.17)	(2.05)
Gov Trans to HH	-1.402***	-1.736***	-1.838***
	(-4.17)	(-6.18)	(-6.45)
Gini	0.347	0.351	
	(1.80)	(1.77)	
Spending R&D	-12.58		
	(-1.69)		
Constant	-11.77	-18.97	-3.497
	(-0.53)	(-0.84)	(-0.16)
Observations	34	34	34
R^2	0.827	0.808	0.787
Adjusted R^2	0.788	0.774	0.758
AIC	257.4	258.8	260.5
BIC	268.1	268.0	268.1

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.2: Ordinary least squares regression estimates for the best subset model, and its backward stepwise selection derivatives

Note: In order to ensure consistency in model comparisons, throughout this section regressions have been limited to observations shared by all possible explanatory variables. Therefore, coefficients may slightly differ with a model where all possible observations are included.

Source: IFs version 7.14

Table A.2 shows the regression outputs of these three models. Though Model 1 appears to be the highest quality according to the AIC and BIC measures, the variables measuring inequality (Gini) and spending on R&D are not significant at the 95% confidence level, after removing spending on R&D (the variable with the lowest t-statistic) we are left with Model 2. While Model 2 reports a higher R^2, the variable measuring spending on R&D continues to be insignificant and so is removed from the model. The resulting Model 3 is comprised of four variables, education, business regulation, tax rate, and government transfers to households, each with significant coefficients and with the expected sign.

The model passes Ramsey-RESET test for omitted variables with a p-value = 0.87. This is further supported by the Shapiro-Wilk test (p = 0.10), which confirms that the model's residuals are normally distributed. According to the Breusch-Pagan / Cook-Weisberg test (p = 0.36) and White test (p = 0.93) heteroskedasticity is not present in the model. Furthermore, by testing the variance inflation factors of explanatory variables we find that the model does not appear to exhibit multicollinearity.

Informal labor residing inside the informal sector

Using a similar process as applied to the specification of the informal labor share model, we began with the universe of variables identified above, remove those already used to forecast informal labor share, and identified a best subset based on relative quality as indicated by AIC and BIC measures using a best sub-set technique.[43] Table A.3 reports the log likelihood (LL), Akaike's information criterion (AIC), and the Bayesian information criterion (BIC) to assess the best regressions for each predictor sub-set.

# of IVs	LL	AIC	BIC
1	-87.01	178.02	180.38
2	-84.05	174.10	177.64
3	-81.24	**170.49**	**175.20**
4	-81.12	172.25	178.14
5	-81.12	174.25	181.31

Table A.3: Best subset results for the informal labor share inside the informal sector model. The best models identified by the AIC and BIC values have been bolded. Models specific to each of these sub-sets are listed in Footnote 43.
Source: IFs version 7.14

The resulting best sub-set model (Model 1 in Table A.4) contains variables that are not significant at the 95% confidence level. Therefore we performed a backward stepwise selection, removing the least significant variables from the model until all remaining variables were found to be significant, determining that informal labor share alone specified the most comparatively robust and significant model (Model 6).

[43] The rows in Table A.1 correspond to the following variable subsets: (1) Informal Labor; (2) Informal Labor, Corruption; (3) Informal Labor, Corruption, Labor Growth; (4) Informal Labor, Corruption, Rule of Law, Labor Growth; (5) Informal Labor, Corruption, Rule of Law, Unemployment, Labor Growth.

	Model 1	Model 2	Model 3
Informal Labor	0.464***	0.470***	0.420***
	(5.16)	(5.37)	(4.93)
Corruption	3.021	2.706	
	(1.70)	(1.72)	
Rule of Law	-10.94		
	(-0.41)		
Constant	37.23*	32.77***	45.58***
	(2.64)	(3.72)	(9.37)
Observations	31	31	31
R^2	0.511	0.508	0.456
Adjusted R^2	0.456	0.472	0.437
AIC	231.5	229.7	230.8
BIC	237.2	234.0	233.6

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.4: Ordinary least squares regression estimates for the best subset model, and its backward stepwise selection derivatives
Source: IFs version 7.14

The model passes Ramsey-RESET test for omitted variables with a p-value = 0.78. This is further supported by the Shapiro-Wilk test (p = 0.10), which confirms that the model's residuals are normally distributed. According to the Breusch-Pagan / Cook-Weisberg test (p = 0.88) and White test (p = 0.42) heteroskedasticity is not present in the model.

Informal GDP share (NOE formulation) and shadow economy

Since informal economic production is predicated on informal factors of productivity, namely labor, we consider the informal employment variable a core explanatory variable in our models of informal GDP share (NOE formulation) and the shadow economy. With this variable fixed in our models, we began specification of the informal economy model by running a backwards-stepwise regression, eliminating non-significant variables based on their t-scores. The surviving model (Model 4) is also found to have the highest relative quality as indicated by AIC and BIC measures.

	Model 1	Model 2	Model 3	Model 4
Informal Labor	0.172	0.146	0.193***	0.190***
	(1.83)	(1.59)	(3.93)	(4.17)
Spending R&D	-2.566*	-2.817*	-2.860*	-2.808*
	(-1.99)	(-2.21)	(-2.25)	(-2.28)
Corruption	-1.569	-1.679	-1.593	-1.458*
	(-1.47)	(-1.58)	(-1.52)	(-2.06)
Rule of Law	1.038	3.145	2.620	
	(0.07)	(0.21)	(0.17)	
Informal Labor Split	0.0972	0.109		
	(0.54)	(0.61)		
Labor Growth	-64.08			
	(-1.25)			
Constant	14.14	12.86	17.73*	18.68***
	(1.32)	(1.20)	(2.51)	(4.24)
Observations	113	113	113	113
R^2	0.549	0.542	0.541	0.541
Adjusted R^2	0.524	0.521	0.524	0.528
AIC	823.2	822.8	821.2	819.3
BIC	842.3	839.2	834.9	830.2

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5: Ordinary least squares regression estimates for all informal economy (NOE formulation) models included in the backward-stepwise regression analysis.
Source: IFs version 7.14

The model passes Ramsey-RESET test for omitted variables with a p-value = 0.65. While the Shapiro-Wilk test indicates that the residuals are not normally distributed (p = 0.02), this result is not necessarily problematic since this criteria "is only required for valid hypothesis testing, that is, the normality assumption assures that the p-values for the t-tests and F-test will be valid. Normality is not required in order to obtain unbiased estimates of the regression coefficients" (Statistical Consulting Group, no date). Furthermore, an interquartile range test on the model's residuals finds no severe outliers. Both an IM-test (p = 0.53) and White's test (p = 0.37) indicate that heteroskedasticity is not present in the model.

An analysis using the shadow economy measure from Elgin and Oztuanli (2012) as the dependent variable reveals a very similar model (Model 4). The surviving model is also found to have the highest relative quality as indicated by the BIC measure.

	Model 1	Model 2	Model 3	Model 4
Informal Labor	0.0939	0.128*	0.133**	0.106*
	(0.99)	(2.44)	(2.68)	(2.28)
Spending R&D	-2.753*	-2.780*	-2.865*	-3.122*
	(-2.14)	(-2.17)	(-2.32)	(-2.54)
Corruption	-1.611	-1.547	-1.749*	-1.740*
	(-1.52)	(-1.48)	(-2.46)	(-2.44)
Labor Growth	-77.14	-78.31	-76.83	
	(-1.51)	(-1.54)	(-1.52)	
Rule of Law	-3.591	-3.969		
	(-0.24)	(-0.26)		
Informal Labor Split	0.0782			
	(0.44)			
Constant	33.71**	37.17***	35.70***	35.68***
	(3.15)	(5.18)	(7.99)	(7.93)
Observations	112	112	112	112
R^2	0.487	0.486	0.486	0.475
Adjusted R^2	0.458	0.462	0.467	0.460
AIC	815.3	813.5	811.6	812.0
BIC	834.3	829.8	825.2	822.9

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.6: Ordinary least squares regression estimates for all shadow economy models included in the backward-stepwise regression analysis.

Source: IFs version 7.14

The model passes Ramsey-RESET test for omitted variables with a p-value = 0.71. The Shapiro-Wilk test indicates that the residuals are normally distributed (p = 0.10). Furthermore, an interquartile range test on the model's residuals finds no severe outliers. Both an IM-test (p = 0.16) and White's test (p = 0.37) indicate that heteroskedasticity is not present in the model.

	Ramsey-RESET	Shapiro-Wilk	White
Informal Labor	0.87	0.10	0.93
Informal Labor Split	0.78	0.10	0.42
Informal GDP (NOE)	0.65	0.02	0.11
Informal GDP (Shadow)	0.71	0.07	0.37

Table A.7: P-values of the primary validation tests for each of the four final models of informality used in the IFs system.

Appendix A5: Informal Economy Historical Datasets Added to the IFs Model

Table A5.1. Informal economy historical datasets added to the IFs model

Variable	Name in International Futures	Description	Years Covered	Countries Included	Source
Informal Economy as % of GDP, Blended	GDPInformal%Adjustment	Informal economy as percentage of GDP	1995-2010	40	UNECE
Informal Economy as % of GDP, UNECE	GDPInformal%Blended	Informal economy as percentage of GDP	1960-2010	165	Blended (UNECE, Schneider, and Elgin)
Informal (Shadow) Economy as % of GDP, Elgin	GDPInformal%Elgin	Informal economy as percentage of GDP, used to initialize shadow economy variable	1950-2009	160	Elgin and Oztunali (2012)
Informal Economy as % of GDP, Schneider	GDPInformal%Schneider	Informal economy as percentage of GDP	1999-2007	161	Schneider and Enste (2012)
% of Total Female Labor in Informal Construction Sector	LaborInformal%ConstructionFemale	Female informal construction labor as a share of total female labor	2004-2006,2008-2010	41	ILO-WIEGO
% of Total Male Labor in Informal Construction Sector	LaborInformal%ConstructionMale	Male informal construction labor as a share of total male labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Labor in Informal Construction Sector	LaborInformal%ConstructionTotal	Informal construction labor as a share of total labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Female Labor in Informal Manufacturing Sector	LaborInformal%ManufacturingFemale	Female informal manufacturing labor as a share of total female labor	2004-2006,2008-2010	42	ILO-WIEGO

% of Total Male Labor in Informal Manufacturing Sector	LaborInformal%ManufacturingMale	Male informal manufacturing labor as a share of total male labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Labor in Informal Manufacturing Sector	LaborInformal%ManufacturingTotal	Informal manufacturing labor as a share of total labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Female Labor in Informal Service Sector	LaborInformal%ServicesFemale	Female informal services labor as a share of total female labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Male Labor in Informal Service Sector	LaborInformal%ServicesMale	Male informal services labor as a share of total male labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Labor in Informal Service Sector	LaborInformal%ServicesTotal	Informal services labor as a share of total labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Labor in Informal Sectors	LaborInformal%TotalAllBlended	Informal labor as a share of total labor	1990-2011	41	Blended (ILO-WIEGO and World Bank)
% of Total Labor in Informal Sectors, ILO-WIEGO	LaborInformal%TotalAllILOWIEGO	Informal labor as a share of total labor	2004-2006,2008-2010	35	ILO-WIEGO
% of Total Labor in Informal Sectors, World Bank	LaborInformal%TotalAllWorldBank	Informal labor as a share of total labor	1990-2011	33	World Bank
% of Informal Labor in Total Labor	LaborInformal%TotalEmp	Informal employment as a share of total employment	2004-2006,2008-2010	35	ILO-WIEGO
% of Total Female Labor in Informal Labor, ILO-WIEGO	LaborInformal%TotalFemale	Female informal labor as a share of total female labor	2004-2006,2008-2010	35	ILO-WIEGO
% of Total Female Labor in Informal Labor, World Bank	LaborInformal%TotalFemaleWB	Female informal labor as a share of total female labor	1990-2013	39	World Bank Gender Statistics

% of Total Male Labor in Informal Labor, ILO-WIEGO	LaborInformal%TotalMale	Male informal labor as a share of total male labor	2004-2006,2008-2010	35	ILO-WIEGO
% of Total Male Labor in Informal Labor, World Bank	LaborInformal%TotalMaleWB	Male informal labor as a share of total male labor	1990-2013	39	World Bank Gender Statistics
% of Total Female Labor in Informal Trade Labor	LaborInformal%TradeFemale	Female informal trade labor as a share of total female labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Male Labor in Informal Trade Labor	LaborInformal%TradeMale	Male informal trade labor as a share of total male labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Labor in Informal Trade Labor	LaborInformal%TradeTotal	Informal trade labor as a share of total labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Female Labor in Informal Transportation Labor	LaborInformal%TransportationFemale	Female informal transportation labor as a share of total female labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Male Labor in Informal Transportation Labor	LaborInformal%TransportationMale	Male informal transportation labor as a share of total male labor	2004-2006,2008-2010	42	ILO-WIEGO
% of Total Labor in Informal Transportation Labor	LaborInformal%TransportationTotal	Informal transportation labor as a share of total labor	2004-2006,2008-2010	42	ILO-WIEGO
Number of Informal Jobs	LaborInformalAll	Number of informal jobs	2004-2006,2008-2010	35	ILO-WIEGO
Number of Informal Jobs Belonging to Females	LaborInformalFemale	Number of informal jobs belonging to females	2004-2006,2008-2010	35	ILO-WIEGO
Number of Informal Jobs Belonging to Males	LaborInformalMale	Number of informal jobs belonging to males	2004-2006,2008-2010	35	ILO-WIEGO

% Female Employment of Total Informal Employment	LaborInformalFemale%AllProdInformal	Female informal employment as a share of total informal employment	2004-2006,2008-2010	35	ILO-WIEGO
% Female Employment of Informal Formal Sector Employment	LaborInformalFemale%FormalSec	Female informal employment in the formal sector as a share of total informal employment in the formal sector	2004-2006,2008-2010	33	ILO-WIEGO
% Female Employment of Informal Household Sector Employment	LaborInformalFemale%HHInformal	Female informal household employment as a share of total informal employment in the household sector	2004-2006,2008-2010	29	ILO-WIEGO
% Female Employment of Informal Sector Employment	LaborInformalFemale%InformalSec	Female informal employment as a share of total informal sector employment	2004-2006,2008-2010	33	ILO-WIEGO
% Formal Sector Employment of Total Informal Employment	LaborInformalFormSec%TotalInformal	Informal employment in the formal sector as a share of total informal employment	2004-2006,2008-2010	33	ILO-WIEGO
% Household Sector Employment of Total Informal Employment	LaborInformalHH%TotalInformal	Informal employment in the household sector as a share of total informal employment	2004-2006,2008-2010	33	ILO-WIEGO
% Informal Sector Employment of Total Informal Employment	LaborInformalInfSec%TotalInformal	Informal employment in the informal sector as a share of total informal employment	2004-2006,2008-2010	33	ILO-WIEGO
% Male Employment of Total Informal Employment	LaborInformalMale%AllProdInformal	Male informal employment as a share of total informal employment	2004-2006,2008-2010	35	ILO-WIEGO
% Male Employment of Informal Formal Sector Employment	LaborInformalMale%FormalSec	Male informal employment in the formal sector as a share of total informal employment in the formal sector	2004-2006,2008-2010	33	ILO-WIEGO
% Male Employment of Informal Household Sector Employment	LaborInformalMale%HHInformal	Male informal employment in the household sector as a share of total informal household sector employment	2004-2006,2008-2010	29	ILO-WIEGO
% Male Employment of Informal Sector Employment	LaborInformalMale%InformalSec	Male informal employment as a share of total informal sector employment	2004-2006,2008-2010	33	ILO-WIEGO

Source: IFs version 7.14