Consumption Taxes, Redistribution and Informality

Pierre Bachas, Lucie Gadenne & Anders Jensen*

Abstract

We study how the presence of large informal sectors in developing countries impacts the distributional properties of consumption taxes. We assemble a dataset of household expenditure using micro-data from 20 countries at different levels of economic development. Using the place of purchase to proxy for informal consumption, we show a large negative relation between informal consumption shares and households' total expenditure, which is robust to product and geography controls. This implies that consumption taxes are de-facto progressive: households in the top decile pay 70% more taxes as a share of expenditure than households in the bottom decile. Finally, we build a model of optimal commodity taxation in the presence of informal consumption, which we calibrate to our data. We find that optimal tax rates are less differentiated across products with an informal sector. Tax exempting necessities, such as food, is rarely optimal as it leads to only a marginal gain in progressivity.

JEL Classification: H21, H23, H32, O23

*Pierre Bachas: World Bank Research, pbachas@worldbank.org. Lucie Gadenne: University of Warwick, Institute for Fiscal Studies and CEPR, l.gadenne@warwick.ac.uk. Anders Jensen: Harvard Kennedy School and NBER, Anders_Jensen@hks.harvard.edu. The findings and conclusions are those of the authors and do not represent the views of the World Bank. We thank Elie Gerschel, Mariana Racimo, Roxanne Rahnama and Alvaro Zuniga for excellent research assistance. Financial support from the World Bank Development Economics' Research Support Budget is gratefully acknowledged.

1 Introduction

Large informal sectors constrain developing countries' capacity to levy direct income taxes (Besley and Persson, 2013; Kleven et al., 2016; Jensen, 2019). As a result, these countries rely on consumption taxes to raise revenue, which are easier to enforce (Gordon and Li, 2009) but considered regressive (Warren, 2008). The negative distributional impact of consumption taxes features prominently in academic and policy debates on the role of tax policy to curb inequality (Lustig, 2018). These debates however do not take into account that in developing countries a large share of household consumption occurs in the informal - untaxed - sector. While there is theoretical work on tax policy in the presence of informal workers and firms (Boadway and Sato, 2009; Gadenne, 2018), little is known about informal consumption and its implication for optimal tax policy.

In this paper we study the distributional consequences of consumption taxes in the presence of informal consumption. We assemble a new household-level dataset using expenditure surveys for 20 countries spanning the development spectrum. Using the place of purchase to proxy for informal consumption (e.g. street stalls vs supermarkets), we find a large negative correlation between informal consumption shares and total household expenditure. This implies that the effective tax rate of households in the top decile is 70% more taxes, as a share of expenditure, than households in the bottom decile. To characterize how the presence of informal consumption impacts both equity and efficiency, we then construct a model of optimal commodity taxation. Calibrating the model to our data, we find that with an informal sector optimal tax rates are less differentiated across products. In particular exempting food products from taxes, a policy often undertaken by developing countries, is unlikely to be optimal as it only leads to a marginal increase in progressivity.

We assemble a dataset of household expenditure by place of purchase to derive new stylized facts on informal consumption across the income distribution. A major constraint in studying informality is that informal sector purchases are by definition hard to observe and impossible to link to a consumer's income level. Our innovation is to assign to each place of purchase a likelihood of remitting taxes: this enables us to proxy for the share of household consumption purchased from the informal sector by product category. The key intuition behind our approach is that large modern retailers are much more likely to be formal than small stores and markets (Kleven et al., 2016). At a broad level, the place of purchase of expenditures can be classified into four categories: (1) non-market consumption (e.g. self-production), (2) non brick-and-mortar retailers (e.g.

street stalls), (3) small brick-and-mortar retailers (e.g. corner shops) and (4) large retailers (e.g. supermarkets). We always assign categories (1) and (2) to the informal sector, category (4) to the formal sector, and consider different assumptions for category (3).[1]

Our first contribution is to describe informal consumption patterns within and across countries. Based on our classification, 70 to 80% of expenditure is informal in Sub Saharan Africa countries, as compared to only 20% in the richest Latin American countries. We then characterize the "Informality Engel Curve" (IEC): the relationship between the share of expenditure spent in the informal sector and household income. We find that the IEC slopes are downward-sloping in all countries in our sample: poor households always spend a higher share of their budget on informal products than rich households. The relationship between the IEC slopes and countries' GDP per capita follows an inverse U-shaped pattern, with particularly steep IECs in middle income countries. In these countries the informal expenditure share falls by 9 percentage points when household income doubles. We then investigate potential explanations for the negative correlation between income and informal consumption. We find that geography plays a role - rural households consume significantly more from the informal sector than urban households, on average - but that controlling for narrow location only reduces the IEC slopes by a quarter. Product composition matters more in explaining the slopes: controlling for detailed (4-digit level) product type reduces the slopes by 45%. IEC slopes remain economically and statistically significant, even once we control jointly for location and product types.

Our second contribution is to characterize the implications of these differences in informal expenditure shares across households for optimal tax policy. Assuming a full pass-through of taxes onto prices, we quantify the implications for the distributional properties of commodity sales taxes. We find that commodity taxes are progressive even if a uniform tax is levied on all products: because IECs are downward sloping, the budget share that the richest decile pays in taxes is nearly twice that paid by the poorest decile on average across all the countries in our sample. Once informal consumption is taken into account exempting food products from taxation, a policy often undertaken by developing countries, only modestly increases the overall progressivity of the tax system.

These results imply that taking the informal sector into account makes consumption taxes more attractive from an equity perspective. There is, however, an extra efficiency

[1]Our baseline assumption splits category (3) into: (a) corner shops, which tend to be small and owner-operated, are assigned to the informal sector, and (b) specialized shops, which are more heterogeneous in size, are assigned to the formal sector.

cost of using commodity taxes in the presence of an informal sector, insofar as commodity taxes distort the relative price of informal and formal sector products. We formalize this equity-efficiency trade-off by constructing a multi-person Ramsey model of commodity taxation to consider optimal tax policy in the presence of an informal sector (formally, we introduce an informal sector in the model of Diamond, 1975). This model yields two key insights. First, optimal commodity tax rates will be higher in the presence of an informal sector as long as i) IECs are downward sloping and ii) the elasticity of substitution between formal and informal varieties is small relative to the own price elasticity. Second, taking into account the informal sector leads to less variation in optimal rates across commodities if commodities which are consumed more by the poor also have steeper downward-sloping IECs than commodities consumed more by the rich. Intuitively, setting lower rates on goods consumed more by the poor is less effective for redistribution if the poor mostly purchase these goods from the informal sector.

Calibration of the model using the patterns of informal expenditure shares observed in our data indicate that optimal tax policy does in practice look different once the informal sector is taken into account. We find that informality-adjusted optimal tax rates are higher than those obtained when there is no informal sector, particularly for food, alcohol & tobacco, and fuel & utilities, all goods that are consumed more by the poor. Rates on these products are higher when adjusting for the informal sector as long as the ratio of elasticity of substitution to own-price elasticity is lower than 0.4.

1.1 Related Literature

This paper contributes directly to the literature which studies how the informal sector impacts policy design. Previous work focuses on the implications of the 'supply side' of informality, by modelling and measuring incentives associated with informality either at the firm-level (DeSoto, 1989; De Paula and Scheinkman, 2010; La Porta and Shleifer, 2014), the worker-level (Gerard and Gonzaga, 2016; Kumler et al., 2013; Jensen, 2019), or both (Ulyssea, 2018). Our paper complements this literature by considering the 'demand side' of informality (consumers' choice of formal or informal retailers), and provides a novel method to measure the informal sector by using the place of purchase information available in many household surveys. Our results imply that the enforcement policies considered by this literature, which reduce the size of the informal sector, may shift part of the burden of taxation towards poorer households and thus have adverse distributional consequences.

This paper is similarly related to the literature on optimal tax design under imperfect enforcement (Allingham and Sandmo, 1972; Cremer and Gahvari, 1993, Best et al., 2015; Boadway and Sato, 2009; Gordon and Li, 2009) which mostly focuses on the efficiency properties of taxation. Our main innovation with respect to this literature is to introduce equity considerations by studying optimal commodity taxation under imperfect enforcement with heterogeneous households (see also Gadenne, 2018). This allows us to characterize how the existence of informal retailers affects the equity-efficiency trade-offs associated with tax policy. This paper also speaks to the literature on the optimal tax mix between direct and indirect taxes over the course of economic development (Burgess and Stern, 1993; Huang and Rios, 2016): our findings imply that tax systems in developing countries can be progressive even in the absence of direct income taxes.

Our results also contribute to the smaller literature on the redistributive properties of consumption taxes in developing countries. Recent empirical studies consider the progressivity of overall tax systems in developing countries (Lustig et al., 2012;Lustig, 2018) or the redistributive impact of consumption tax policies (Harris et al., 2018; Abramovsky et al., 2015). These studies do not take into account how informal expenditure shares vary across households and conclude that consumption taxes at best do not affect inequality. In contrast, we find that consumption taxes are strongly progressive once informal sector consumption is accounted for. Our methodological approach is related to Jenkins et al. (2006) and Muñoz and Cho (2003) who assess the redistributive consequences of tax reforms in, respectively, the Dominican Republic and Ethiopia, using retailer information to classify some purchases as informal. Relative to these papers, our analysis is based on a more recent and larger set of surveys which include internationally comparable sets of expenditure items and places of purchases. This database allows us to provide novel facts on informal consumption patterns across development, which we relate to both progressivity assessments of current consumption tax systems and calibration of optimal tax rates for all countries in our sample.

Finally, this paper adds to the wider literature on consumption choices and development. Prior studies have focused on documenting how expenditure shares on specific goods vary with income, including the well-established Engel curve for food (Deaton and Paxson, 1998; Anker et al., 2011; Pritchett and Spivack, 2013). We study instead how expenditure shares spent on informal products vary both within countries and across levels of development - the Informality Engel Curve. Our main empirical result - a negative correlation between income and informal expenditure share - holds both across and

within narrow groups of products, and in all countries considered. A related set of studies have focused on how the availability of different retailer types affect consumption choices accross the household income distribution. Atkin et al. (2018b) find that modern retail entry in Mexico leads to substitution across store-types, with a larger elasticity for wealthier households. This heterogeneity is consistent with a household valuation of product quality that is increasing in income (Faber and Fally, 2017), and with the adoption of costly durables (like cars) that are complementary to modern retail shopping (Lagakos, 2016). Both these mechanisms can similarly explain the patterns of informal consumption we observe. Finally, our methodology is related to papers that study macroeconomic changes over development using internationally comparable micro-data (Bick et al., 2018; Jensen, 2019).

The rest of the paper is structured as follows. Section 2 describes our data sources and the methodology to measure informal consumption and assign places of purchase to a formality status. Section 3 provides new stylized facts across development on consumption patterns. Section 4 quantitatively assesses the implications of the observed IEC for progressivity of a consumption tax. Section 5 derives the optimal commodity tax rule in the presence of informal consumption, and Section 6 calibrates the optimal tax rates using the observed IECs in all countries in our sample. Section 7 concludes.

2 Data

2.1 Sample of Countries

Our dataset combines 20 nationally representative household expenditure surveys. We retain surveys which satisfy four criteria: (1) it is nationally representative and from the 21st century, (2) it asks about the place of purchase for each item in its expenditure modules, (3) the expenditure modules are structured as open consumption diaries rather than pre-filled diaries for specific products, and (4) the answer to the place of purchase question is rarely missing, in particular for product categories food, clothing and household goods.[2]

Table 1 lists alphabetically the 20 countries in the sample, with survey names, years, number of households, and number of expenditure items per household. Countries in the sample are principally located in Latin America and Sub-Saharan Africa, with the

[2]Criteria (5) is obtaining data access: for a few countries we identified surveys which appeared to satisfy all criteria but could not obtain the micro-data (most prominently Egypt and Turkey).

Table 1: Household Expenditure Surveys

Country	Survey	Year	GDP pc	# of HHs	Items/HH
Brazil	POF	2008	15,483	56,049	48.0
Burkina Faso	EICVM	2010	671	8,404	45.6
Burundi	ECVM	2014	320	6,681	89.8
Cameroon	ECAM	2014	1,447	10,303	95.8
Chile	EPF	2017	24,085	15,237	129.2
Colombia	ENIG	2007	9,711	42,733	31.8
Costa Rica	ENIGH	2014	15,140	5,705	66.5
Dem. Rep. Congo	E123	2005	474	12,098	15.5
Ecuador	ENIGHUR	2012	5,969	41,760	109.6
Mexico	ENIGH	2014	18,149	19,459	56.8
Morocco	ENCDM	2001	8,217	14,243	89.9
Mozambique	IOF	2009	857	10,659	26.7
Niger	ENCBM	2007	378	3,980	61.1
Papua New Guinea	HIES	2010	2,517	3,811	14.4
Peru	ENAHO	2017	13,434	43,530	56.4
Rep. of Congo	ECOM	2005	458	5,002	55.0
Rwanda	EICV	2014	703	14,419	89.5
South Africa	IES	2011	13,498	25,325	44.0
Tanzania	HBS	2012	2,946	10,168	187.0
Uruguay	ENGIH	2006	16,246	7,043	77.5

This tables lists alphabetically the countries in this study. GDP per capita is in PPP USD from World Bank WDI. Items per households correspond to the average number of expenditure items per household in the diaries.

exceptions of Morocco and Papua New Guinea. The absence of Asian countries is due to the structure of their expenditure surveys, where the place of purchase is rarely recorded (criteria 2 not satisfied), and which have many missing answers when asked (criteria 4 not satisfied).[3] Appendix A.1 details the data sources and the surveys considered for inclusion, and explains the reasons for discarding specific surveys (table A3).

Table A1 provides details on the structure of the surveys' expenditure modules for the countries in the sample. While all surveys satisfy criteria (1)-(4) above, their structure varies across countries. In particular, we note that surveys have different numbers of expenditure modules, which typically correspond to different time frames of purchases (e.g. separate modules for weekly purchases vs. infrequent quarterly purchases).

[3]Survey design appears to be strongly correlated across countries within regions, showing the influence of regional development partners and/or historical ties across statistical administrations. For example, surveys from francophone countries in sub-Saharan Africa more frequently feature the place of purchase question compared to SSA anglophone countries.

2.2 Assignment of Places of Purchase to Formality Status

We proxy for informality of consumption by focusing on the formality status of the place of purchase recorded in the household diary for each expenditure. In this subsection, we outline the assignment of formality status across different categories of place of purchase. The level of detail and the names of places of purchases necessarily differ across countries. In the assignment process, our aim is to limit the number of country-specific choices: appendix A.2 provides an in-depth description of these choices.

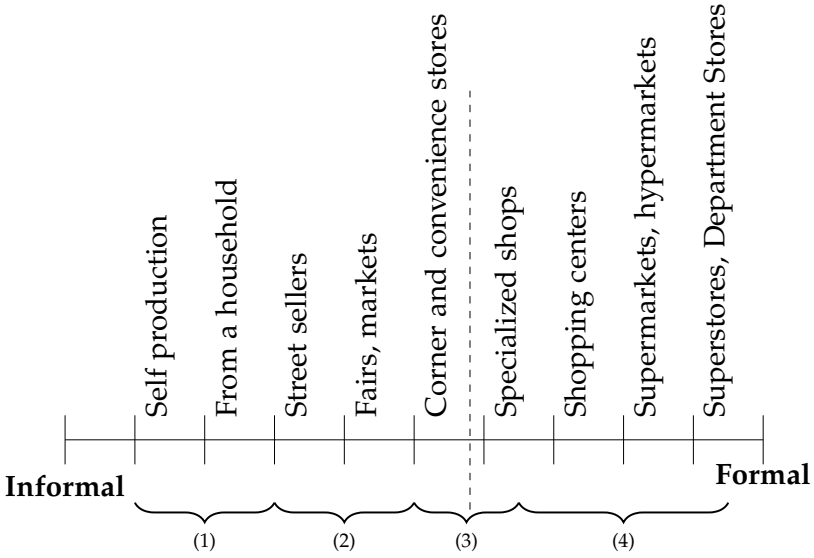
At a broad level, the place of purchase for the bulk of recorded expenditures can be classified into one of four categories: (1) Non-market consumption (e.g. self-production, barter), (2) Market consumption in non brick-and-mortar stores (e.g. street stalls, public markets), (3) Market consumption in small brick-and-mortar stores (e.g. corner shops, specialized stores) and (4) Market consumption in large brick-and-mortar stores (e.g. supermarkets, department stores).

Using this general classification, we assign to *informal* all places of purchases in categories (1) Non-market and (2) Market non brick-and-mortar. Non-market expenditures do not enter the market, and as such are not subject to commodity tax. Purchases in category (2) rarely reflect commodity taxes in final consumer prices. This could first arise because of legal exemptions: in many countries, firms and traders characterized by small revenue are legally relieved from remitting commodity taxes to the government. Non brick-and-mortar stores are very likely to have low enough revenues that they are exempt from legal remittance of commodity taxes. The informality of goods in (2) could also arise de facto: even if liable to remit, transactions in these stores are hard for the tax authority to observe, and therefore to enforce. Furthermore, the absence of a physical location means the enforcement cost of trying to levy taxes on these stores will often outweigh the potential recovered revenue.[4] In contrast, expenditure in large stores (Category 4) is assigned to *formal*. Owing to their revenue size, large stores are almost always legally liable to remit commodity taxes. De facto compliance in these stores is sustained because evasion becomes prohibitively costly when coordination on under-reporting involves too many different agents. The third category, market consumption from small brick-and-mortar stores (3), is arguably the category with the most variation in true formality status across settings. This category includes corner stores which tend to be small and owner-operated, and specialized shops which are more heterogeneous in size. In

[4]The legal decision to exempt small firms from commodity taxes is often in itself motivated by the large administrative costs required for enforcement and the expected compliance rate.

our central scenario, we therefore split Category 3: we assign convenience and corner stores to *informal*, while specialized stores are assigned to *formal*. Figure 1 depicts our suggested formality assignment status across the most commonly used names for places of purchases in our surveys. In the central scenario, expenditures in places of purchase to the left (right) of the dotted line are considered informal (formal).

Figure 1: Assignment of Places of Purchase to Formality



While places of purchases in categories (1)-(4) characterize a large share of expenditure for the countries in our sample, some places of purchase do not fit in this nomenclature and are assigned to category (5) "other places of purchases". The most common places of purchase in category (5) are public institutions, health and educational institutions, online shopping, and entertainment. These places of purchase are likely to be formal and taxable.[5] We assign these places of purchase to *formal*, unless there is a direct indication in the name of a place of purchase that it is relatively more (in)formal than another place of purchase in the same survey. Per example, within health purchases, we assign formal to "public health institution" and informal to "traditional medicine"; within entertainment, we assign formal to "restaurant" and informal to "canteen-truck."

Finally, while a direct criteria for including a survey in our sample is that for household products (food, clothing, appliances) the place of purchase is rarely unspecified, this can be frequent for utilities, education and health product categories. When the

[5]Whether these places of purchases are taxed in practice, is a different matter. For example taxing large online retailers is not technologically difficult, but countries have only slowly adjusted their tax legislation, in part due to difficulties in assessing the jurisdiction where taxes should be remitted.

place of purchase for these product groups is unspecified or "Others", we classify these as formal. Table A5 shows for each country the original names of the places of purchase, our formality assignment, and their share of expenditure. Finally, Appendix A.2 discusses our treatment of place of purchase which do not fall neatly into the above suggested classification (e.g. online, abroad).

2.3 Discussion of Informality Assignment

The idea behind our formality assignment is that places of purchase labelled as formal are expected to generate substantial trails, which can be followed by the tax administration. Formal places of purchase are much more likely to be large in size, both in terms of employees and of sales. Kleven et al. (2016) show theoretically that collusion between firm managers and employees to evade taxes is increasingly costly when the collusive agreement has to be reached with many employees. Similarly, formal stores are likely to have a larger number of customers. Naritomi (2018) shows that if customers have either intrinsic motives or are financially rewarded to report evasion, the likelihood of detected evasion will increase in the size of the customer base. This occurs both through a monitoring effect, and through the per-customer collusion cost the firm must incur to evade sales taxes. Naritomi (2018) also shows that an increase in third party information trails on final consumer sales leads to a larger positive revenue impact in smaller firms (measured by sales), suggesting information trails led larger firms to be more compliant at baseline. She also provides evidence suggesting that collusion costs are important to sustain compliance in stores with larger customer base.

In work in progress, we try to provide direct evidence of the formality status of places of purchases in our sample. We proceed in the following manner: (1) For all countries in our sample, we try to access the census of establishments when it includes the retail sector, and keep firms in the retail sector; (2) We identify potential formality measures: these always include registration statuses and sometimes directly include tax payments (and in particular consumption taxes); and (3) We then show the correlation between formality measures and either directly type of retailers when available, or indirectly with employee size.

Our formality assignment has focused on the tax-status of the place of purchase. We have abstracted from the possibility that informal firms in practice may pay consumption taxes on the purchase of inputs from formal firms. In turn, if these input costs are passed onto final product prices, then customers in informal stores will de facto pay

taxes on some of their goods. Formality in input purchases will therefore lead us to overstate the informal share income-gradient. Recent evidence points to substantial market segmentation between formal and informal firms, so informal firms may purchase little from formal suppliers (Gadenne et al., 2019). We discuss further the implications of modelling taxes paid in inputs in the model (Section 6.3).

3 Informal Consumption Along the Expenditure Distribution

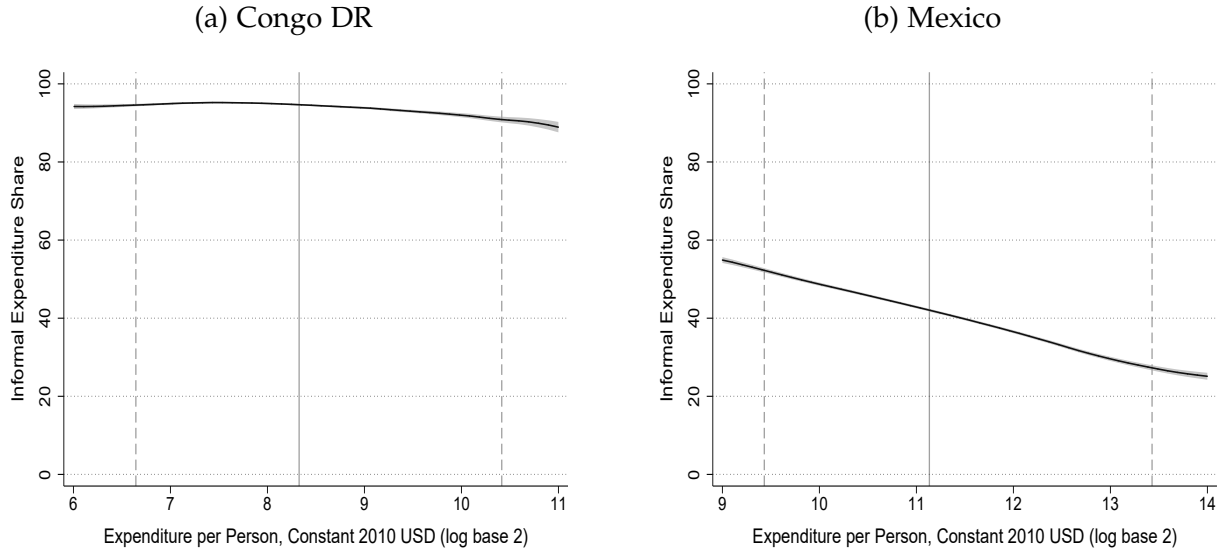
In this section, we use the newly constructed micro-data of expenditures by place of purchase to document new facts on informal consumption, and its relation with income within and across countries. We show how this relation is impacted by the inclusion of controls for household characteristics, location and product codes. Finally, we provide a discussion for what could explain the un-accounted portion of the slopes.

3.1 Informality Engel Curves

Our main object of interest is the relation between the share of informal expenditure and per capita total expenditure of households, which we term the informality Engel curve (IEC). We follow the literature on consumption in developing countries and associate expenditure shares with total expenditure rather than income as in (Deaton and Paxson, 1998; Atkin et al., 2018a). Our formulation is closely related to the typical Engel curve in individual goods (Working, 1943), which plots budget shares on individual goods items against log total household expenditure. As an example, Figure 2 shows non-parametric informality Engel curves for two of the most populous countries in our sample: DR Congo and Mexico. The x-axis is the per capita expenditure of households in 2010 constant USD, measured in log base 2, such that a one unit increase on the horizontal axis corresponds to a doubling of household expenditure. The solid grey line corresponds to the median of each country's expenditure distribution, while the dotted lines correspond to the 5^{th} and 95^{th} percentiles. We use kernel-weighted local polynomials and the shaded areas show the 95% confidence intervals. For example in Mexico, the poorest households consume 55% informally, the median households 40% and the richest households less than 25%. The Engel curves for each individual country are displayed in Figure C.1.

To summarize across countries the information contained in all the informal Engel curves, we focus on two key moments: the country's average informal expenditure share across its population and the slope of its IEC. Taking the slope as a summary measure

Figure 2: Informality Engel Curves



Local polynomial fit of the informality Engel curves for DR Congo and Mexico (All countries in Figure C.1). The informality Engel curves are defined as the share of informal consumption over the per capita expenditure of households. Expenditure is measured in log base 2, such that a one unit increase on the horizontal axis corresponds to a doubling of household's expenditure. The shaded area around the polynomial fit corresponds to the 95% confidence interval. The solid grey line corresponds to the median of each country's expenditure distribution, while the dotted lines correspond to the 5th and 95th percentiles.

assumes a linear relation between informal expenditure share and the log (base 2) of total household expenditure. While the literature has found that Engel curves for specific individual goods are non-linear (Banks et al., 1997; Atkin et al., 2018a), in our setting, visual inspection suggests that linearity is a reasonable approximation for IECs in most countries. The slopes are estimated from the following regression:

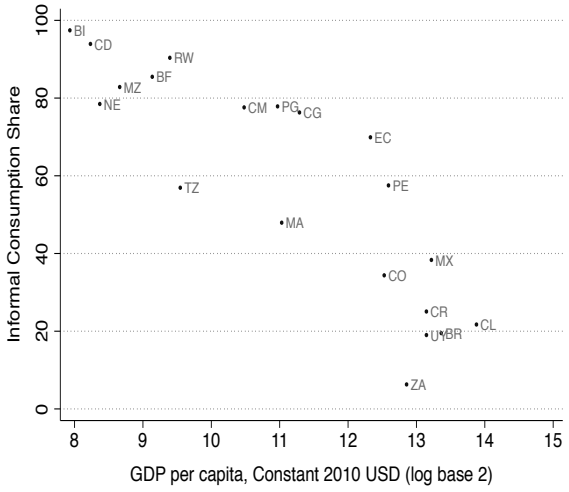
$$Share\ Informal_i = \beta * ln(expenditure\ pc)_i + \varepsilon_i \tag{1}$$

Figure 3, Panel (a), plots the average informal expenditure (as a share of total expenditure) on a country's per capita GDP (2010 constant USD). Based on our classification, 58% of expenditure is informal in the countries in our sample. A country's informal expenditure is highly correlated with its per capita GDP: 70 to 80% of exependiture is informal in Sub Saharan Africa countries, as compared to only 20% in the richest countries in the sample. Panel (b) shows the IEC slopes on GDP per capita. The average slope across countries is 5.4: this implies that as a household doubles its expenditure, its informal expenditure share is reduced by 5.4%. The figure also suggests an inverted-U shaped pattern of IEC slopes across development. In the poorest countries, the slopes are lower than the average: most consumption is informal and only households at the

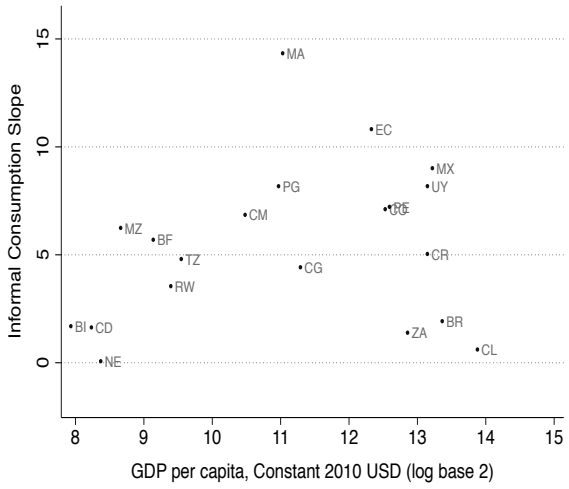
(very) top of the expenditure distribution consume less informally (as exemplified by DR Congo in Figure 2). Similarly, for some of the richer countries, most consumption is already formal and the slopes are relatively flat. Countries of medium income levels display the largest IEC slopes.

Figure 3: Informal Consumption Across Countries

(a) Informal Consumption Share



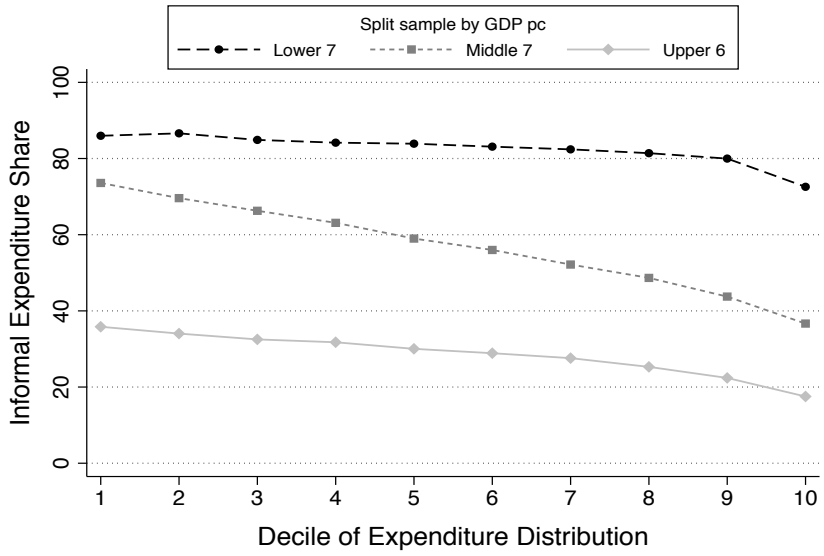
(b) Slope of Informal Engel Curve



Panel (a) shows average informal consumption (as a share of total consumption) in each country on its per capita GDP. Panel (b) shows the within country slope of informal consumption with expenditure on GDP per capita in constant USD: the slope measures the drop in informal consumption for a doubling of households' expenditure, within country. GDP per capita is in constant 2010 USD, and transformed in log base 2 (Source: World Bank, WDI)

In the discussion above we assumed linearity - instead Figure 4 summarizes the distributional statistics of informal expenditure: it shows the average informal consumption for each decile of the within country expenditure distribution, taking the average across groups of countries based on their income levels (Low income, Middle Income, Upper-middle income). On average across all countries, the bottom decile consumes 65% informally while the top decile consumes 40% informally. These results mask substantial heterogeneity by country's income levels (Mirroring the results on the IEC slopes): for the lowest income countries the informal expenditure share is fairly flat, until the top decile where it drops by 10%. For middle and upper-middle income countries the results are starker: bottom decile households consume twice as much informally as compared to the top decile households.

Figure 4: Informal Consumption Distribution



The figure shows the share of informal consumption per expenditure decile grouping countries in three income levels. We take the average across countries within a similar income group, of the countries' share of informal consumption in each expenditure decile. Low-income countries are the 7 poorest countries in the sample (GDP pc below 1,500 USD in PPP) middle-income are the 7 countries with GDP pc between 1,500 and 7,000, and upper-middle income are the 6 countries with GDP pc above 7,000 USD.

3.2 Explaining the IEC slopes: Household Characteristics, Geography & Products

3.2.1 Household Characteristics

In this section we investigate in a regression framework the extent to which the slopes of the informal Engel curves can be explained by observables. We start by considering household characteristics: age, education level, and gender of household head, as well as number of household members. The latter controls for economies of scale in per capita consumption across households of different size (Deaton and Paxson, 1998). Upon inclusion of these control variables, we estimate:

$$Share\ Informal_i = \beta.ln(expenditure\ pc)_i + \Gamma X_i + \varepsilon_i \tag{2}$$

We report results of the regressions in Table 2. As expected, household characteristics have a limited impact on the slopes of the IEC, and on average slighlty increase them: the average IEC slope across countries increases from 5.4 (col1) to 5.8 (col2). Table C1 shows the individual slopes of each country. In all subsequent specifications we maintain these household characteristics as control variables.

Table 2: Average Informal Consumption Slopes

Specification:	**Main**		**Geography**		**Product Codes**			**All**
Avg. of 20 Countries	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Slope	5.6	5.9	4.9	4.3	3.8	3.7	3.1	2.4
Confidence Interval	[5.3,5.8]	[5.6,6.2]	[4.6,5.2]	[4.0,4.7]	[3.5,4.0]	[3.5,4.0]	[2.9,3.3]	[2.2,2.7]
# of p-values < 0.05	19	19	19	18	18	18	18	18
R squared	0.14	0.16	0.20	0.38	0.45	0.47	0.48	0.51
Household Characteristics		X	X	X	X	X	X	X
Urban/Rural			X					
Geo. Districts				X				X
COICOP 2-dig					X			
COICOP 3-dig						X		
COICOP 4-dig							X	X

This table shows the average slope of the Informal Engel curve across countries for different specifications. The slopes are estimated from: $Share\,Informal_i = \beta.ln(expenditure\,pc)_i + \Gamma X_i + \varepsilon_i$, where the dependent variable is the informal expenditure share and the explanatory variable is the log expenditure pp. Controls include household characteristics (household size, age, gender, and education of head), geographic indicators (urban/rural and districts), and product codes at the 2nd, 3rd and 4th level of the United Nation's COICOP classification. While all countries follow the COICOP at the 2nd level (12 categories: food, clothing etc.), Brazil and Peru have specific product classifications at lower levels. The geographic control "district" refers to the lowest geographical level available in each survey.

3.2.2 Geography

One hypothesis is that large modern stores are only located in dense urban centers, and that the IEC slope is driven mainly by differences in income between rural and urban households. If this is the case, then controlling for density or urbanization would shrink the slopes. All surveys come with an urban -rural indicator, and we construct a "district" measure which corresponds to the lowest geographical level available in the survey.[6] In Table 2 we successively control for an urban-rural dummy (col3) and for geographical district dummies (col4). The average slopes fall by 17% when controlling for rural-urban and by 21% when controlling for geographical districts. This points to a limited role for access and density in explaining the informal Engel curves. The differences in the levels of informal expenditure between rural and urban households are interesting in and

[6]The density and definition of districts differs across countries. In the future we aim to construct a more comparable measure.

of themselves: as one would expect, rural households consume 13% more informally (64% versus 51% of expenditure share). Figure C.4 shows the difference in rural and urban informal expenditure by country. The slopes within urban and rural are not very different, and actually slightly higher for urban.

3.2.3 Product Composition

Another hypothesis is that rich households consume different products than poor households, and that the products consumed by the rich are only sold in formal stores (e.g. televisions) and vice-versa. Then controlling for goods type would eliminate the IEC slopes. We test for this hypothesis by controlling for product codes at different levels of classification. For 17 surveys we have the United Nations's COICOP product classification, while for Brazil, Morocco and Peru we directly use the countries' product classification.[7] At the 2 digit level there are 12 categories, which correspond to food, clothing, furnishings etc. The 3 digit contains 47 products and the 4 digit 117 products.[8] We then estimate the slopes of the informality Engel curve, controlling for product composition by running the following regressions:

$$Share\ Informal_{ip} = \beta.ln(income\ pp)_i + \alpha_p + \Gamma X_i + \varepsilon_{ip} \tag{3}$$

Where $Share\ Informal_{ip}$ is the share of informal expenditure by household i on product p and α_p are product fixed effects. We weight for inverse probability of inclusion of household and for importance of the product in total expenditure.

Table 2 reports the results when controlling for product codes at the 2nd, 3rd and 4th level of COICOP (Columns 5, 6, 7). Controlling for large product groups (COICOP-2) explains 37% of the slopes. Further product codes only explain slightly more of the slopes, and even at the 4th digit of COICOP over half of the IEC slopes remain unexplained. Figure C.3 displays visually these results for each country by showing the change in slopes when adding sequentially one more level of product codes. We note in particular that the slope is reduced the most for countries in the middle of the income distribution, flattening the relation between IEC slopes and countries' per capita GDP. Finally, we note that simultaneous controls for products and geography only add little to the product controls (C1)

[7] For these countries we only have crosswalks for COICOP2 but not at lower levels. This implies that their slopes are not directly comparable to other countries' when controlling for COICOP 3 and 4.

[8] Figure C.3 provides an example of a COICOP branch going from the 2nd to the 4th digit.

3.3 Further investigation of choice of place of purchase

In this sub-section, we leverage additional data contained in a sub-sample of surveys to further investigate the choice of formal versus informal stores. In four countries (Morocco, Dem. Republic of Congo, Rep. of Congo, Burundi), the expenditure module asked customers to pick the main reason for choosing a particular place of purchase for each purchase. Respondents could pick from a list of answers, which we group into five categories: access, price, quality, store attributes, and other. Access is defined as a combination of necessity and proximity, while store attributes includes quality of reception, offering of credit, and homogeneity of products on display. In turn, we study the frequency of different categories that are picked as the most important criteria, separately for formal and informal stores.

Two main patterns emerge from table 3. First, between roughly a third (Dem Rep Congo, Rep. of Congo) and a half (Burundi, Morocco) of the respondents choose access as the main reason for choice of store. Large shares of households in this sub-sample of countries are likely to have too low income to invest in costly durables, such as cars, which reduce access constraints (Lagakos, 2016). The large substitution across types of store observed in the Mexican setting when a modern retailer enters a local market (Atkin et al., 2018b) is also consistent with constrained access. In contrast, Allcott et al. (2017) find that the entry of a new supermarket leads to no substitution across types of stores in the US - a setting where car ownership is widespread. The second pattern is a price-quality trade-off in formal versus informal stores. In particular, respondents cite quality as more important than price when choosing formal stores, while in informal stores price is cited as more important than quality. The reversal of price vs quality rankings between formal and informal stores is observed in all four countries. This price-quality trade-off is quantitatively important: with the exception of Morocco, price (quality) is cited as the main reason for choosing an informal (formal) store for at least a third of households in all samples. Combined with our main finding that richer households shop in formal stores, this quality-price trade-off is consistent with recent findings by Faber and Fally (2017). In the U.S. context, the authors show that richer households value quality more than poorer households and stores (endogenously) cater to wealthy households by producing higher quality products. Because of economies of scale in quality production, firms that cater to wealthy households are larger in size than those that sell to poor households. Since larger stores are more likely to be formal (cf Section 2.3), this mechanism can give rise to the observed IECs.

Table 3: Main Reason for Choosing Place of Purchase

	Morocco			Dem. Rep. Congo			Rep. of Congo			Burundi		
Reason	Total	Formal	Informal	Total	Formal	Informal	Total	Formal	Informal	Total	Formal	Informal
Access	56%	54%	56%	27%	24%	28%	35%	41%	34%	50%	41%	50%
Price	24%	7.3%	27%	35%	24%	36%	32%	20%	36%	28%	22%	29%
Quality	7.8%	18%	6.2%	19%	43%	18%	16%	28%	14%	8.5%	32%	6.9%
PoP Attributes	7.8%	9.6%	7.5%	10%	4.6%	10%	8.8%	5.7%	9.7%	4.2%	1.0%	4.4%
Other	4.6%	11%	3.7%	8.8%	4.2%	7.9%	8.3%	5.7%	7.5%	9.4%	4.9%	9.7%

In these four countries, "Access" is defined as a combination of proximity and necessity and "Attributes of PoP" is defined as a combination of homogeneity of products, offering of credit, and quality of reception.

4 Implications for Tax Progressivity

This section considers what the patterns of informal sector consumption depicted above imply for the redistributive potential of commodity sales taxes. In this section, we abstract from behavioral responses and efficiency considerations, and assume that the pass-through of taxes to prices is one, therefore showing the mechanical implications of our results. To do so, we consider the share that each household pays in taxes under four different scenarios with different tax policies and assumptions about the environment. The two types of policies we consider are i) a uniform tax on sales of all goods and ii) a tax exemption of all food products, with a uniform tax on all other goods. This second policy approximates the type of policy often used by developing countries to try to make their tax systems more progressive. Using our data we simulate the share of each household's budget spent on taxes under both policies for two different assumptions: first we assume that taxes are paid on all products, regardless of the retailer type, then we assume that taxes are paid only by retailers we classify as belonging to the formal sector. We set tax rates under all four scenarios such that the total amount of tax revenues raised is equal to 10% of total consumption, and assume throughout that taxed firms fully pass-through the tax to consumer prices.

Figure 5 presents the average taxed budget share for each decile under these four scenarios.[9] Figure A on the left plots the average taxed budget share under the assumption that the government can potentially tax all products (no informal sector), for a uniform tax policy (black line) and for a tax policy that exempts food products from taxation and levies a uniform tax on all other products (grey dashed line). All households pay 10% of their budget in tax under the uniform tax scenario, by assumption. Exempting food

[9] Each country is given equal weight in the average.

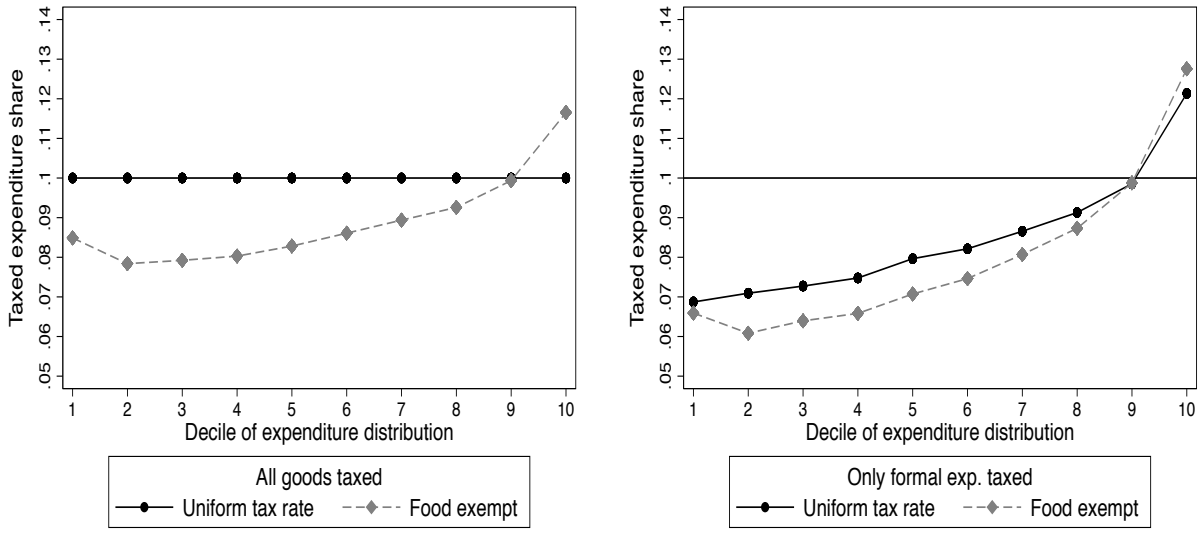
products from taxation makes commodity taxes slightly progressive: the poorest three deciles pay roughly 8% of their budget in tax once food is exempt from taxation, this shortfall in tax revenues compared to the uniform tax scenario is compensated for by a higher tax on non-food products, paid for only by the richest decile, which pays 11.6% of its budget in taxes.

Figure B on the left of Figure 5 plots the average taxed budget share under both policy scenarios under the assumption that only formal retailers pay taxes. We see that the patterns of informal sector consumption described above imply that even a uniform commodity tax enables the government to redistribute: the poorest households pay 6-7% of their budget in taxes on average, compared to 12% for the richest households (black line). Moreover, all households below the 80th percentile pay on average less than 10% of their budget in tax, despite the fact that this tax scenario by assumption levies 10% of total consumption in tax: most of the burden of taxation is borne by the top two deciles. Exempting food from taxation (dashed grey line) still makes the commodity tax system more progressive, but only barely: the taxed budget share of the poorest decile falls from 6.4 to 6.1, that of the richest decile increases from 12.1 to 12.6 when food is exempt from taxation. Results for each country separately, available from the authors upon request, show that the redistributive dividend due to the informal sector is highly correlated with the slope of the informality Engel curves depicted above. The steeper the informality Engel curve, the more taking into account the informal sector makes simulated taxed budget shares increase with income.

Overall, these results indicate that taking into account how patterns of informal sector consumption vary with income makes commodity sales taxes more progressive than typically assumed in the literature. The informal sector thus makes commodity sales taxes more attractive from an equity perspective. There is however an extra efficiency cost of using commodity taxes in the presence of an informal sector, relative to a world with no informal sector, insofar as commodity taxes distort the relative price of informal and formal sector products. The next section explicitly considers both these equity and efficiency considerations by looking at how the existence of an informal sector affects optimal tax policy.

Figure 5: Taxed Expenditure Shares under Four Scenarios

(a) All Products Taxed (b) Only formal expenditures taxed



The figure shows the share of households' budget paid in taxes under four different scenarios. Each point is on average across all the countries in our sample. Graph A assumes that all products can be taxed by the government, graph B assumes that only products sold by formal retailers can be taxed. For both graphs the dark line depicts taxed budget shares for a uniform tax levied on all goods and the grey dashed line depicts taxed budget shares for a scenario in which food products are exempt from taxation. The government's budget constraint is held constant over all scenarios, such that total tax revenues raised are always equal to 10% of total consumption.

5 Optimal Taxation with an Informal Sector

The following section considers how the existence of an informal sector affects optimal commodity taxation. To do so we introduce the possibility that some goods cannot be taxed because they are produced in the informal sector in the multi-person Ramsey model of commodity taxation of Diamond (1975).

5.1 Set-up

There is a continuum of mass one of households i which different incomes y^i. Households have homogeneous preferences over J different commodity, and for each commodity j over two varieties $j0$ and $j1$, where variety $j0$ is potentially produced in the informal sector. To simplify the exposition of the results we assume no cross-price elasticity between the different j commodities but a positive cross-price elasticity ϵ_{j01} between varieties: we assume formal and informal varieties of each commodity are substitutes. These assumptions are relaxed in the Appendix.

The government sets different tax rates on each commodity type j, but cannot tax

different varieties differently. The consumer price of variety $l = 0, 1$ of good j, which we call variety jl, is thus $p_{jl} = q_{jl}(1 + t_j)$ if variety jl is produced in the formal sector, and $p_{jl} = q_{jl}$ if it is produced in the informal sector, where the producer price q_{jl} is exogenous. Consumer maximization yields household demands for the goods $x^i_{jl}(p, y^i)$, where p is the vector of consumer prices. We write $v(p, y^i)$ indirect utility of household i and s^i_j the budget share that household i spends on commodity j.

5.2 Optimal taxation when all goods can be taxed

We start by considering a world with no informal sector: one in which both varieties of each commodity are taxed. The government chooses the tax rates t_j to maximize:

$$W = \int_i G(v(p, y^i))di + \mu \sum_j t_j(q_{j0}x_{j0} + q_{j1}x_{j1}) \tag{4}$$

where $x_{jl} = \int_i x^i_{jl}(p, y^i)di$ is total demand for variety jl, μ is the marginal value of the public good and $G()$ is the governments' social welfare function, increasing and concave. We write g^i household i's social marginal welfare weight which represents how much the government values giving an extra unit of income to household i, and g the average social marginal welfare weight (see Saez and Santcheva, 2016).[10] We assume throughout that g^i is non-increasing with household income. We often refer below to the case $\mu = g$ to simplify expressions, this corresponds to a government that has no preference for taxing revenues away from households unless it enables redistribution.

Welfare maximization yields the following expression for optimal tax rates $\tau^*_j = \frac{t^*_j}{1-t^*_j}$:

$$\tau^*_j = \frac{(\mu - g) - \int_i (g^i - g)\beta^i s^i_j / s_j di}{-\mu \epsilon_j} \tag{5}$$

where $\beta^i = \frac{y^i}{\int_i y^i di}$ is the share of household i's income in total income and $s_j = \int_i s^i_j di$ the average budget share spent on j. We write ϵ_j the uncompensated price elasticity of demand for commodity j; it is a weighted sum of the own price elasticities of demand for varieties $j0$ and $j1$ and the cross-variety price elasticities:

$$\epsilon_j = \alpha_{j1}(\epsilon_{j1} + \epsilon_{j1,j0}) + (1 - \alpha_{j1})(\epsilon_{j0} + \epsilon_{j0,j1}) \tag{6}$$

[10] Formally $g^i = \frac{\partial G(v(p,y^i))}{\partial v(p,y^i)} \frac{\partial v(p,y^i)}{\partial y^i}$.

where $\alpha_{j1} = x_{j1}/x_j$ is the share of variety $j1$ in total consumption of j.[11]The more the two varieties of good j are substitutes, the lower the price elasticity of demand for good j (in absolute value) relative to price elasticity of variety jl.

This expression simply states that optimal tax rates are increasing with the government's relative preference for public revenues $(\mu - g)$, lower for commodities that are consumed more by the poor (such that $Cov(g^i, s^i)$ is positive), and lower for commodities that are very price elastic.

5.3 Optimal taxation in the presence of an informal sector

We now consider optimal commodity tax rates in a world in which variety $j0$ for each commodity j is produced in the informal sector and therefore untaxed. The government now maximizes

$$W = \int_i G(v(p, y^i))di + \mu \sum_j t_j q_{j1} x_{j1} \tag{7}$$

Welfare maximization yields the following expression for the optimal tax rates:

$$\tau_j^{**} = \frac{(\mu - g) - \int_i (g^i - g)\beta^i s_{j1}^i / s_{j1}}{-\mu \epsilon_{j1}} \tag{8}$$

Setting $\mu = g$ we obtain the following expression for the ratio of optimal tax rates under both scenarios:

$$\frac{\tau_j^*}{\tau_j^{**}} = \frac{\epsilon_{j1}}{\epsilon_j} \frac{\int_i (g^i - g)\beta^i s_j^i / s_j}{\int_i (g^i - g)\beta^i s_{j1}^i / s_{j1}} \tag{9}$$

Consider first the case of a commodity which all households equally consume from the informal sector, such that $s_j^i/s_j = s_{j1}^i/s_{j1}$: a commodity with a flat informality Engel curve. The presence of an informal sector does not impact the extent of redistribution achievable by taxing this commodity: the right-hand-side of (9) becomes $\frac{\epsilon_{j1}}{\epsilon_j}$, so that differences between the optimal rates are only a function of households' behavioral responses. Assuming that own-price elasticities are similar for the formal and informal varieties this implies that the optimal rate on commodity j is lower in the presence of an informal sector. Intuitively this is because in the world with no informal sector t_j effectively taxes two substitute goods $j1$ and $j0$ in the same way, whereas with an informal

[11]Using Slutsky symmetry this can be rewritten as function of the compensated cross price elasticity between the commodities and the income elasticities η_{j0} and η_{j1}: $\epsilon_j = \alpha_{j1}\epsilon_{j1} + (1 - \alpha_{j1})\epsilon_{j0} + \epsilon_{0,1}^C - (\alpha_1 \eta_1 s_0 + (1 - \alpha_1)\eta_0 s_1)$.

sector setting a non null tax t_j distorts the relative prices of these two goods, leading to an additional efficiency cost of taxation that's increasing in the cross-price elasticity.

To abstract from such efficiency considerations consider now a commodity for which the cross-variety price elasticity is zero: there is no substitution between the formal and the informal sector and therefore no difference in efficiency cost between the two tax rates. The ratio in (9) is now less than 1 as long as s_j^i/s_{j1}^i is decreasing with income: as long as the informality Engel curve is downward sloping. Intuitively when the poor consume a higher share of the commodity from the informal sector the burden of taxation is borne more by richer households when the informal sector cannot be taxed. Equity considerations thus imply that optimal commodity rates will be higher when there is an informal sector as long as informality Engel curves are downward-sloping.

Combining equity and efficiency considerations yields the following proposition.

Proposition 1. ***Taking into account the informal sector increases the optimal tax rate for a good as long as i) the informality Engel curves is downward sloping for this good and ii) the price elasticity of substitution between formal and informal varieties is not too large compared to the good's own price elasticity.***
Formally, $\frac{\tau_j^*}{\tau_j^{**}} \leq 1$ *as long as the cross price elasticity is small relative to the own price elasticity,* $\epsilon_{j1,j0} \leq -\epsilon_{j1}(1 - \frac{\int_i (g^i - g)\beta^i s_j^i / s_j}{\int_i (g^i - g)\beta^i s_{j1}^i / s_{j1}})$. *This condition can only hold if* $\frac{\int_i (g^i - g)\beta^i s_j^i / s_j}{\int_i (g^i - g)\beta^i s_{j1}^i / s_{j1}} < 1$, *ie if informality Engel curves are downward sloping.*[12]

At the limit, note that for some commodities introducing an informal sector can even change the sign of the optimal taxrate. Consider an inferior commodity whose optimal rate without an informal sector is negative ($Cov(g^i, \beta^i s^i)$ negative). If the poor are a lot more likely to consume it from the informal sector than the rich, such that $Cov(g^i, \beta^i s_{j1}^i)$ is positive, the optimal rate in the presence of an informal sector is positive.

Overall expression (9) indicates that there is an equity-efficiency trade-off when considering how the presence of an informal sector affects optimal commodity taxes. Efficiency concerns imply that optimal taxes should be lower in the presence of an informal sector but equity considerations tend to increase them.

Finally, we can study how the informal sector affects the desirability of rate differentiation across products by looking at the extent to which optimal rates vary across commodities j. To focus on the role of rate differentation in redistribution, assume price

[12] This proposition is derived under the assumptions $\epsilon_{j1} = \epsilon_{j0}$ and $\epsilon_{j1,j0} = \epsilon_{j0,j1}$. The latter assumption is equivalent to assuming there are no income effects.

elasticities do not vary across goods. The level of optimal rate differentation is then set by the variation in $Cov(y^i, s^i_j)$ across goods j when all goods are taxed, and the variation in $Cov(y^i, s^i_{j1})$ when only the formal sector is taxed. This leads to the proposition:

Proposition 2. ***Taking into account the informal sector leads to less variation in optimal rates across goods if goods with downward sloping Engel curves also have steeper downward sloping informality Engel curves than goods with upward sloping Engel curves.***
To see this note that a downward sloping informality Engel curve decreases the value of $Cov(y^i, s^i_{j1})$ *compared to* $Cov(y^i, s^i_j)$ *for commodities for which* $Cov(y^i, s^i_j) < 0$ *(downward sloping Engel curves), but it increases the value of* $Cov(y^i, s^i_{j1})$ *compared to* $Cov(y^i, s^i_j)$ *for commodities for which* $Cov(y^i, s^i_j) > 0$ *(upward sloping Engel curves) .*

6 Calibrated Optimal Tax Rates

This section calibrates optimal commodity sales tax rates defined in expressions (5) and (8), using the patterns of informal consumption by income level presented in section 3.

6.1 Calibrated government and household preferences

Our baseline calibrations assumes that the government is poverty minimizing, and therefore places a higher weight on households in the first three deciles. We set the marginal social welfare weight g equal to 2 for households in these deciles, and equal to 1 for all other households. The marginal value of public funds, μ, is set equal to the average marginal social welfare weight. We set the own-price elasticity to -0.7 for all goods at baseline, a value within the range discussed by Deaton (1997). The parameters β_i (share of household i's income in total income) is obtained from our data, using total consumption as a proxy for income, and so are the parameters s^i_j (budget share spent by i on good j) and s^i_{j1} (budget share spent by i on the formal variety of good j).Finally, the model predicts that the relative value of optimal rates with and without an informal sector will be a function of the magnitude of the cross-variety price elasticity relative to the own price elasticity. We consider different values for this cross-variety price elasticity below.

We consider optimal rates for the 12 COICOP-2 digits product categories: food, alcohol and tobacco, clothing and footwear, housing and utilities, household equipment, health, transport, communication, recreation and culture, education, restaurants and hotels, miscellaneous goods and services.

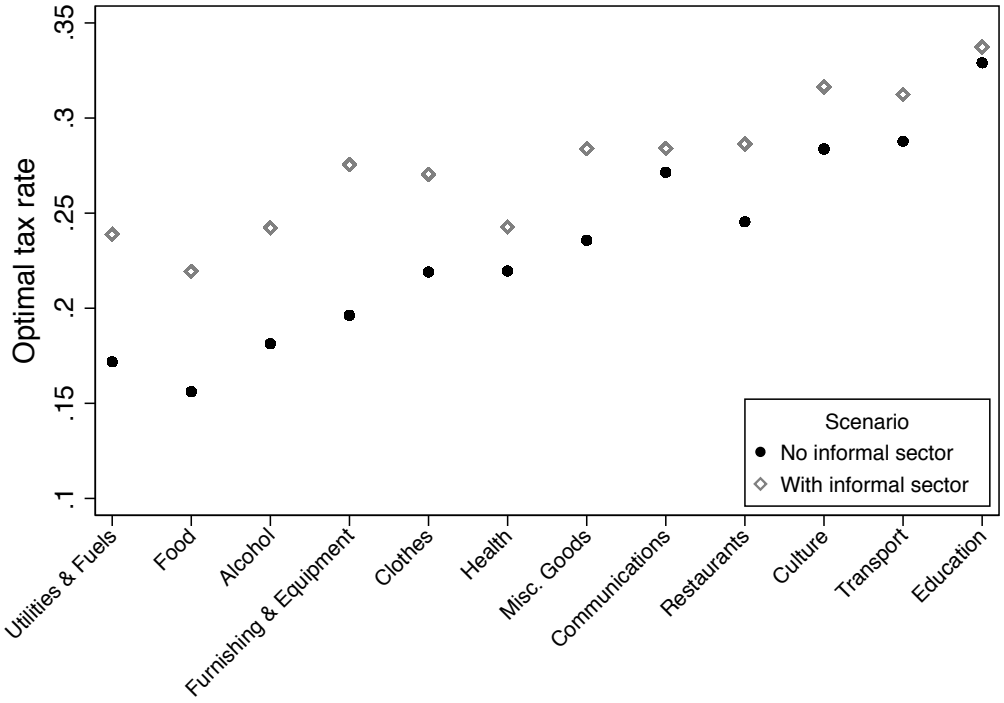
6.2 Results

We start by considering results assuming no substitution between varieties ($\epsilon_{j1,j0} = 0$). This assumption ensures that the efficiency cost of using commodity taxes is the same regardless of whether the informal sector is taken into account, so differences between optimal tax rates can only be due to equity considerations. Results are presented in Figure 6 and Table 4. The black dots depict the average optimal rates obtained under the assumption that all goods can be taxed. We see they are always positive, reflecting the fact that richer households consume more of each product category, but vary substantially across products: food and housing are taxed the least with a 15% rate, whilst education and culture are taxed at over twice that rate. These different tax rates reflect differences in Engel curves for these products, the redistributive potential of taxing food at a lower rate in the absence of an informal sector is well illustrated in Figure 5.

The grey diamonds depict the average optimal rates obtained under the assumption that goods sold by informal retailers cannot be taxed. Comparing these with the optimal tax rates obtained without taking the informal sector into account yields two key conclusions. First, as predicted by the model, these optimal rates are always at least as high as those obtained not taking into account the informal sector. On average the difference between the two types of optimal rates is small (3 percentage points), but it is larger for products with steep informality Engel curves: food, alcohol,housing equipment and housing and utilities. There is barely any difference on the other hand for products that are rarely purchased from informal retailers, such as education, health, communication or transports. Second, there is less variation in optimal rates across goods once the informal sector is taken into account. The highest optimal rate is 2.3 times the lowest rate when there is no informal sector, but this ratio falls to 1.7 when the informal sector is taken into account. This reflects the fact that a large share of the differences across products in the correlation between budget shares and income, which drive the differences across rates, can be explained by differences in the correlation between informal budget shares and income. The Engel curve for food, for example, is less steep if we consider only food consumption from formal retailers than if we consider food consumption from all retailers, because the informality Engel curve for food, depicted above, is steeply decreasing with income.

Table 4 also presents average optimal tax rates for different values of the cross-variety price elasticity as a share of the own price elasticity. We see that, as expected, optimal rates obtained under the assumption of no informal sector increase as the cross-variety

Figure 6: Average Optimal Tax Rates, no Efficiency Considerations



The figure shows the average optimal tax rate for each product, where each country is given equal weight in the average. The black dots are optimal rates when we assume all goods can be taxed, the grey diamonds are optimal rates when we assume goods sold by informal retailers cannot be taxed. These rates obtained by assuming that the cross-variety price elasticity is zero, see the text for a description of the other parameters used.

price elasticity increases (optimal rates obtained taking into account the informal sector are unaffected by this parameter, by definition). This is because the higher this elasticity the lower the own price elasticity (in absolute value) that matters in the absence of an informal sector, as explained in the model above. Average optimal rates are higher than in the presence of an informal sector for all products considered once the cross-variety price elasticity reaches 40% of the (absolute) value of the own price elasticity. There is always less variation in optimal rates across product categories once the informal sector is taken into account, regardless of the value taken by the cross elasticity share. This is because the redistributive potential of rate differentiation isn't affected by price elasticities. Future work will discuss how optimal rates vary when different calibration choices are relaxed, in particular the specification of government differences.

Table 4: Calibrated Optimal Tax Rates

	With informal sector	No informal sector		
Cross elasticity share	All	0	0.2	0.4
	(1)	(2)	(3)	(4)
Food	0.20 (0.07)	0.15 (0.03)	0.18 (0.04)	0.22 (0.04)
Housing	0.21 (0.12)	0.15 (0.05)	0.18 (0.06)	0.22 (0.07)
Alcohol	0.24 (0.08)	0.18 (0.03)	0.22 (0.06)	0.27 (0.07)
Household equipment	0.26 (0.08)	0.19 (0.06)	0.23 (0.06)	0.28 (0.07)
Clothes	0.26 (0.08)	0.21 (0.07)	0.25 (0.07)	0.30 (0.08)
Health	0.23 (0.07)	0.22 (0.06)	0.26 (0.07)	0.32 (0.08)
Misc	0.26 (0.07)	0.24 (0.07)	0.27 (0.08)	0.34 (0.09)
Restaurants	0.27 (0.09)	0.25 (0.05)	0.29 (0.05)	0.35 (0.05)
Communication	0.26 (0.10)	0.25 (0.10)	0.30 (0.11)	0.35 (0.12)
Transport	0.28 (0.04)	0.28 (0.04)	0.32 (0.05)	0.39 (0.05)
Culture	0.30 (0.08)	0.28 (0.06)	0.32 (0.07)	0.39 (0.08)
Education	0.33 (0.09)	0.34 (0.10)	0.39 (0.10)	0.45 (0.11)
Mean across products	0.26 (0.03)	0.23 (0.05)	0.27 (0.06)	0.32 (0.06)
SD across products	0.068 (0.017)	0.071 (0.018)	0.078 (0.019)	0.087 (0.02)

This table presents average optimal tax rates for each product category in the first twelve lines, across all products categories in the penultimate line, and the standard deviation across products in the last line. Standard deviations across countries are presented in parentheses. The first column presents optimal rates obtained under the assumption that only varieties sold by informal retailers can be taxed, the value taken for the cross-variety price elasticity does not affect these rates. The last three columns present optimal rates obtained under the assumption that all varieties can be taxed, for three different values of the cross-variety price elasticity relative to the own price elasticity : 0 in column 2, 0.2 in column 3 and 0.4 in column 4.

6.3 Discussion: Supply chain considerations

Our analysis thus far ignores the possibility that informal firms de facto pay commodity taxes because they purchase intermediate inputs from formal firms which themselves charge commodity taxes on their products. In practice however formal firms may partially withhold taxes from their informal clients (Keen, 2008). If informal firms pass-through the taxes they pay on their formal input purchases to the prices of their products, consumers will de facto pay taxes on their purchases from the informal sector. The pass-through of commodity taxes to prices will still be higher in the formal sector than in the informal sector, but the relative redistributive dividend due to the informal sector will be muted relative to that depicted in Figure 5 in which we assume zero pass-

through.[13] There is evidence of substantial market segmentation between firms with different tax status in supply chains (Gadenne et al., 2019), suggesting informal firms may buy little from formal suppliers. In future work we will consider the implications of relaxing our assumption that informal firms do not pay taxes on their inputs, using data on the share of inputs purchased from formal suppliers for a few countries.

7 Conclusion

In this paper we studied the empirical and theoretical implications of informal consumption for the design and progressivity of consumption taxes in developing countries. We find that the informal expenditure share is decreasing in income - a relation we termed the Informality Engel Curve (IEC). Interestingly, while the average levels of informal consumption are largest in low income countries, the income-slopes are steepest in middle-income countries. We show that these consumption patterns substantially increase the progressivity of consumption taxes. We finally derive expressions for optimal commodity taxes in the presence of informal consumption, and calibrate optimal taxes for the 20 countries in our sample using the observe informal consumption patterns.

Our modelling and measurement of demand for informality has been in partial equilibrium and we have ignored supply side responses in terms of changing formality status of stores, as a function of the tax rate. In practice, the changes in informal consumption we find across levels of income will have implications for the equilibrium shares of informal firms and workers in an economy. Our IECs provide a set of moments that can be used to calibrate a general equilibrium model to assess overall welfare impacts of tax and regulatory policies in settings with informality.

References

Abramovsky, L., O. Attanasio, and D. Phillips (2015): "Value Added Tax policy and the case for uniformity: empirical evidence from Mexico," *IFS Working Papers*.

Allcott, H., R. Diamond, J.-P. Dubé, J. Handbury, I. Rahkovsky, and M. Schnell (2017): "Food Deserts and the Causes of Nutritional Inequality," .

Allingham, M. G. and A. Sandmo (1972): "Income tax evasion: a theoretical analysis," *Journal of Public Economics*, 1, 323–338.

[13] In a model with monopolistic competition, the pass-through of taxes to prices equals one in the formal sector and equals the share of input costs purchased from formal suppliers in the informal sector.

Anker, R. et al. (2011): "Engel's law around the world 150 years later," *Political Economy Research Institute Working Paper No*, 247.

Atkin, D., B. Faber, T. Fally, and M. Gonzalez-Navarro (2018a): "A New Engel on the Gains from Trade," Tech. rep., Working paper.

Atkin, D., B. Faber, and M. Gonzalez-Navarro (2018b): "Retail globalization and household welfare: Evidence from mexico," *Journal of Political Economy*, 126, 1–73.

Banks, J., R. Blundell, and A. Lewbel (1997): "Quadratic Engel curves and consumer demand," *Review of Economics and statistics*, 79, 527–539.

Besley, T. and T. Persson (2013): "Chapter 2 - Taxation and Development," in *handbook of public economics, vol. 5*, ed. by M. F. Alan J. Auerbach, Raj Chetty and E. Saez, Elsevier, vol. 5 of *Handbook of Public Economics*, 51 – 110.

Best, M., A. Brockmeyer, H. Kleven, J. Spinnewijn, and M. Waseem (2015): "Production vs Revenue Efficiency With Limited Tax Capacity: Theory and Evidence From Pakistan," *Journal of Political Economy*, 123(6), 1311–1355.

Bick, A., N. Fuchs-Schündeln, and D. Lagakos (2018): "How do hours worked vary with income? Cross-country evidence and implications," *American Economic Review*, 108, 170–99.

Boadway, R. and M. Sato (2009): "Optimal tax design and enforcement with an informal sector," *American Economic Journal: Economic Policy*, 1, 1–27.

Burgess, R. and N. Stern (1993): "Taxation and development," *Journal of economic literature*, 31, 762–830.

Cremer, H. and F. Gahvari (1993): "Tax evasion and optimal commodity taxation," *Journal of Public Economics*, 50, 261–275.

De Paula, A. and J. A. Scheinkman (2010): "Value-added taxes, chain effects, and informality," *American Economic Journal: Macroeconomics*, 2, 195–221.

Deaton, A. (1997): *The analysis of Household surveys: a microeconometric approach to development Policy*, John Hopkins University Press.

Deaton, A. and C. Paxson (1998): "Economies of scale, household size, and the demand for food," *Journal of political economy*, 106, 897–930.

DeSoto, H. (1989): *The other path*, Harper & Row New York.

Diamond, P. A. (1975): "A many-person Ramsey tax rule," *Journal of Public Economics*, 4, 335–342.

Faber, B. and T. Fally (2017): "Firm heterogeneity in consumption baskets: Evidence from home and store scanner data," .

Gadenne, L. (2018): "Can Rationing Increase Welfare? Theory and an Application to India's Ration Shop System," Tech. rep., CEPR Working Paper 13080.

Gadenne, L., T. Nandi, and R. Rathelot (2019): "Taxation and Supplier Networks: Evidence from India," Tech. rep., Mimeo, University of Warwick.

Gerard, F. and G. Gonzaga (2016): "Informal labor and the efficiency cost of social programs: Evidence from the brazilian unemployment insurance program," Tech. rep., National Bureau of Economic Research.

Gordon, R. and W. Li (2009): "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation," *Journal of Public Economics*, 93, 855–866.

Harris, T., D. Phillips, R. Warwick, M. Goldman, J. Jellema, K. Goraus, and G. Inchauste (2018): "Redistribution via VAT and cash transfers: an assessment in four low and middle income countries," Tech. rep., Institute for Fiscal Studies Working Paper W18/11.

Huang, J. and J. Rios (2016): "Optimal tax mix with income tax non-compliance," *Journal of Public Economics*, 144, 52–63.

Jenkins, G. P., H. P. Jenkins, and C. Y. Kuo (2006): "Is the value added tax naturally progressive?" *Available at SSRN 897677*.

Jensen, A. (2019): "Employment Structure and the Rise of the Modern Tax system," Tech. rep., National Bureau of Economic Research.

Keen, M. (2008): "VAT, tariffs, and withholding: Border taxes and informality in developing countries," *Journal of Public Economics*, 92, 1892–1906.

Kleven, H. J., C. T. Kreiner, and E. Saez (2016): "Why can modern governments tax so much? an agency model of firms as fiscal intermediaries," *Economica*, 51.

Kumler, T., E. Verhoogen, and J. A. Frías (2013): "Enlisting employees in improving payroll-tax compliance: Evidence from Mexico," Tech. rep., National Bureau of Economic Research.

La Porta, R. and A. Shleifer (2014): "Informality and development," *Journal of Economic Perspectives*, 28, 109–26.

Lagakos, D. (2016): "Explaining cross-country productivity differences in retail trade," *Journal of Political Economy*, 124, 579–620.

Lustig, N. (2018): *Commitment to Equity Handbook: Estimating the Impact of Fiscal Policy on Inequality and Poverty*, Brookings Institution Press.

Lustig, N., G. Gray Molina, S. Higgins, M. Jaramillo, W. Jiménez, V. C. Paz, C. Pereira, C. Pessino, J. Scott, and E. Yañez (2012): "The Impact of Taxes and

Social Spending on Inequality and Poverty in Argentina, Bolivia, Brazil, Mexico, and Peru: A Synthesis of Results," *Center for Global Development Working Paper*.

Muñoz, S. and S. S.-W. Cho (2003): "Social Impact of a Tax Reform; The Case of Ethiopia," IMF Working Papers 03/232, International Monetary Fund.

Naritomi, J. (2018): "Consumers as tax auditors," *CEPR Discussion Paper*.

Pritchett, L. and M. Spivack (2013): "Estimating Income/Expenditure Differences Across Populations: New Fun with Old Engel's Law," *Center for Global Development Working Paper*.

Saez, E. and S. Santcheva (2016): "Generalized Social Marginal Welfare Weights for Optimal Tax Theory," *American Economic Review*, 106, 24–45.

Ulyssea, G. (2018): "Firms, informality, and development: Theory and evidence from Brazil," *American Economic Review*, 108, 2015–47.

Warren, N. (2008): "A Review of Studies on the Distributional Impact of Consumption Taxes in OECD Countries," Tech. rep., OECD Social, Employment and Migration Working Papers No. 64.

Working, H. (1943): "Statistical laws of family expenditure," *Journal of the American Statistical Association*, 38, 43–56.

A Data Appendix

A.1 Dataset Construction

Inclusion Criteria

The dataset is constructed based on 20 nationally representative household expenditure surveys. We retained surveys which satisfied the following five criteria:

1. The household expenditure survey is nationally representative and dates from the 21st century.
2. The expenditure survey includes a variable on the place of purchase: for each consumption item there is information on where each consumption item was purchased. The place of purchase names should have enough details for reliable classifications into formal or informal sectors, as further outlined in Appendix A.2. In particular, there should be enough detail to classify at least one place of purchase to the formal sector and at least one place of purchase to the informal sector.
3. The expenditure modules in the survey are structured as open consumption diaries, rather than pre-fill diaries for specific products.
4. The place of purchase variable rarely contains missing values; particularly for food, clothing, and household goods product categories.
5. We were able to obtain access to the data.

Table A1 lists the countries in the sample, with summary statistics and the structure of each survey's expenditure modules.

Data Sources and Coverage

We obtained the data principally from two sources: (i) the World Bank Microdata Library and (ii) national statistical agencies, with the exception of South Africa, for which the data came from the University of Cape Town (refer to column 3 of Table A1). The first step for accessing data started with the restricted-access World Bank Microdata Library, where we examined National Household Income and Expenditure, Living Standards, and Budget Surveys to see if criteria (1)-(4) above were satisfied. The datasets that satisfied such criteria ranged in their ease of access: while some countries' survey microdata

Table A1: Household Expenditure Surveys

			HOUSEHOLD SURVEYS						STRUCTURE of EXPENDITURE MODULES					
Country	**Survey-Year**	**Source**	**# HH**	**Avg HH Size**	**Avg Exp/HH** (Current USD)	**# Exp/HH**	**# PoP**	**Urban**	**# Modules**	**Module Freq**	**Durables**	**Self Prod.**	**Comments**	**Product Code**
Brazil	Pesquisa de Orçamentos Familiares 2008	Stat. Office	56,049	3.39	4956.63	48.0	753	0.77	8	Weekly, Monthly Quarterly, Yearly	Included	Included	Food, non-food modules separated	Country-specific
Burkina Faso	L'Enquete integrale sur les Conditions de Vie des Menages 2010	Stat. Office	8,404	6.80	1549.98	45.6	45	0.34	1	Yearly	Included	Included		COICOP
Burundi	Enquete sur les Conditions de Vie des Menages 2014	WB Microdata	6,681	4.89	1730.41	89.8	13	0.31	23	Bi-Weekly, Quarterly Biannual, Yearly	Included	Included	Food, transport, clothing rec., comm. modules separated	COICOP
Cameroon	Quatrieme Enquete Camerounaise Aupres des Menages 2014	WB Microdata	10,303	4.57	5314.52	95.8	17	0.53	1	Daily	Included	Included		COICOP
Chile	Encuesta de Presupuestos Familiares 2017	Stat. Office	15,237	4.00	18423.48	129.2	17	0.52	1	Monthly	Included	Unavailable		COICOP
Colombia	Encuesta Nacional de Ingresos y Gastos 2007	Stat. Office	42,733	3.87	2445.6	31.8	24	0.92	5	Monthly Quarterly, Yearly	Included	Included	Public services, urban, rural modules separated	COICOP
Costa Rica	Encuesta Nacional de Ingresos y Gastos de los Hogares 2014	Stat. Office	5,705	3.38	10451.76	66.5	41	0.66	1	Monthly	Included	Included		COICOP
Dem. Rep. Congo	Enquete 123 sur l'emploi, le secteur informel et les conditions de vie des menages 2005	WB Microdata	12,098	5.44	272.13	15.5	13	0.27	1	Yearly	Included	Included		COICOP
Ecuador	Encuesta Nacional de Ingresos y Gastos de Hogares Urbanos y Rurales 2012	WB Microdata	41,760	3.87	42,332.81	109.6	75	0.74	7	Daily, Weekly, Bi-Weekly Monthly, Quarterly, Biannual, Yearly	Included	Included	Food, non-food modules separated	COICOP
Mexico	Encuesta Nacional de Ingreso-Gasto de los Hogares 2014	WB Microdata	14,243	3.78	4553.84	56.8	19	0.57	1	Quarterly	Included	Included		Country-specific
Morocco	Enquête nationale sur la consommation et les dépenses des ménages 2001	WB Microdata	14,243	6.00	36235.96	89.9	47	0.55	17	Weekly, Monthly, Two-months Quarterly, Yearly	Unavailable	Included	Food, non-food modules separated	Country-specific
Mozambique	Inquérito ao Orçamento Familiar 2009	WB Microdata	10,659	4.75	1360.08	26.7	6	0.50	6	Daily, Monthly, Yearly	Included	Included	Self-consumption, transfer modules separated	COICOP
Niger	Enquête Nationale sur le Budget et la Consommation des Ménages 2007	WB Microdata	3,980	7.17	4298.79	61.1	15	0.48	6	Daily, Weekly, Monthly Quarterly, Biannual, Yearly	Unavailable	Included	Food, non-food modules separated	COICOP
Papua New Guinea	Household Income and Expenditure Survey 2010	WB Microdata	3,811	5.55	5119.05	14.4	6	0.45	1	Bi-Weekly	Included	Included		COICOP
Peru	Encuesta Nacional Hogares 2017	Stat. Office	43,530	3.84	4144.44	56.4	41	0.63	8	Yearly	Included	Included rec., comm. modules separated	Food, transport, clothing,	Country-specific
Rep. of Congo	Enquete Congolaise aupres des Menages pour L'Evaluation de la Pauvrete 2005	WB Microdata	5,002	5.20	865.65	55.0	17	0.79	1	Yearly	Included	Included		COICOP
Rwanda	Integrated Household Living Conditions Survey 2014	WB Microdata	14,419	1.00	768.99	89.5	11	0.16	8	Monthly, Yearly 10 weeks	Included	Included	Food, non-food modules separated	COICOP
South Africa	Income and Expenditure Survey 2011	DataFirst	25,325	3.75	4516.21	44.0	6	0.64	1	Yearly	Unavailable	Unavailable		COICOP
Tanzania	Household Budget Survey 2012	WB Microdata	10,168	4.58	983.04	187.0	13	0.34	2	Monthly	Unavailable	Included	Food, non-food modules separated	COICOP
Uruguay	Encuesta Nacional de Gastos e Ingresos de los Hogares 2006	Stat. Office	7,043	2.92	5536.95	77.5	39	0.65	1	Monthly	Unavailable	Unavailable		COICOP

were openly accessible to the public for download, others were licensed and required applications through the World Bank which would then sometimes contact the respective country's national statistical agency for approval.

If the applicable data was inaccessible through the World Bank Microdata Library, we searched the country-specific statistical agency websites. For some countries, this simply involved downloading the micro-data off of the website, while for others, we made formal data requests. This second step came with varying degrees of success: we obtained data through this route for most Latin American countries, however this process was not always successful elsewhere.

The countries that ultimately satisfied the criteria for inclusion span four regions of the world, with the greatest number of countries concentrated in Sub-Saharan Africa and Latin America and the Caribbean, as detailed in Table A2. Unfortunately we were not able to find any East Asian countries, with a question on the place of purchase in their household expenditure surveys. In Eastern Europe and Central Asia surveys sometimes ask this question, but the number of options for place of purchase are minimal, and the type of stores is often missing, thus not satisfying inclusion criteria 2 and 4.

Table A2: Regional Survey Representation

Region	**# Countries**	**Pop. of Surveyed Countries** (Millions)	**Total Pop.** (Millions)	**Proportion of pop.**
Sub-Saharan Africa	10	320.2	1216	30.0 %
Middle East, North Africa	1	35.74	444.32	8.04 %
Europe & Central Asia	0	0.00	915.55	0.00 %
Latin American & Caribbean	8	462.49	568.13	81.4 %
East Asia & Pacific	1	8.25	2314	0.35 %

While there were a number of surveys that outwardly appeared to satisfy all of the four main selection criteria (namely, Argentina, Belarus, Gambia, Ghana, and Turkey), we were constrained by issues of data access. Table A3 further details countries that were considered for inclusion in our sample, but were ultimately discarded for failing to satisfy any of (1) - (5) from the first page.

Table A3: Discarded Household Expenditure Surveys

Country	Survey	Year	Why Discarded
Argentina	Encuesta Nacional de Gastos de los Hogares	2013	(5) data access constraints
Armenia	Integrated Living Conditions Survey	2016	(4) PoP missing
Belarus	Household Sample Survey	2010	(5) data access constraints
Bosnia & Herzegovina	Household Budget Survey	2007	(3) PoP asked as purchasing habit
Chad	Enquete sur la Consommation et le Secteur Informel	2003	(2) PoP vague
El Salvador	Encuesta de Hogares de Propositos Multiples	2010	(4) PoP missing; limited consumption categories
The Gambia	Integrated Household Survey	2003	(5) data access constraints
Ghana	Living Standards Survey	2006	(5) data access constraints
Guatemala	Encuesta Nacional sobre Condiciones de Vida	2000	(4) PoP missing; limited consumption categories
Mauritius	Household Budget Survey	2012	(4) PoP missing
Montenegro	Household Budget Survey	2009	(4) PoP missing
Nicaragua	Encuesta Nacional de Hogares sobre Medicion de Nivel de Vida	2014	(3) PoP asked as purchasing habit; (5) limited consumption categories
Serbia	Living Standards Measurement Survey	2007	(4) limited consumption categories
Tajikistan	Household Budget Survey	2016	(4) limited consumption categories
Turkey	Household Income and Consumption Expenditures Survey	2009	(5) data access constraints
Ukraine	Household Living Conditions Survey	2012	(5) data access constraints; (4) limited consumption categories

Consumption Module Structure

While each country satisfies (1) to (5), consumption modules are structured differently across countries. Table A1 details their structure and how they can differ - we provide below a summary:

- **Number and frequency of modules**
  - The number of consumption modules range from 1 to 17 modules across countries in the sample. For example, while Costa Rica had one consumption module, Morocco had 17 consumption modules. Modules may also vary based on frequency of expenditures (daily, weekly, monthly, quarterly).
- **Durables**
  - Durable items, which do not have to be purchased frequently (e.g. furniture, motor vehicles, etc.), were included whenever available.
- **Self-production**
  - Self production was included as a "place of purchase" whenever available. In some cases, it was pre-coded in the raw PoP variable, while in others we added it as a PoP based on other variables, such as "mode of acquisition," which included the response of "self/home production."
- **Product codes**
  - Modules have product codes for each consumption item which either follow the official United Nations Classification of Individual Consumption According to Purpose (COICOP) standard or a nationally-specific product classification

scheme - which we harmonized with COICOP codes through a detailed cross-walk.

In spite of the diversity of frequency and number of modules, all surveys are nonetheless structured as open-fill diaries of consumption, in satisfaction of criteria (3) under Section A.1. Figure A.1 obtained from South Africa's 2010-2011 Income and Expenditure Survey serves as an illustrative examples of directions for filling out the household diary, as well as what a typical "open fill" diary of consumption looks like.

INCOME AND EXPENDITURE SURVEY 2010/11
GUIDELINES FOR FILLING IN DAILY ACQUISITIONS

TO BE RECORDED DAILY IN:

Form 1 (pages 6-11)

- All food and non-food items purchased by any member of the household, such as bread, milk, rice, furniture, electric appliances, wood, etc. for the household's consumption as well as to give away as a gift or maintenance.

  Households to record the full price of the item at the time of acquisition.

- All items acquired by the household without paying for them, such as items from own production, (e.g. from own garden or kraal) or from nature (e.g. items from hunting, fishing and gathering of vegetables).
  To be recorded when the item was consumed by the household.

- All items received as gifts or maintenance from someone who is not a household member.

  Households to record the estimated value of the item acquired.

Form 2 (pages 12-13)

- All food and beverages (such as hamburgers, fruit, soft drinks, etc) acquired at restaurants, canteens and other food outlet.
- "Small" acquisitions (such as cigarettes, newspapers, sweets, soft drinks, etc) by individual household members.

  To be recorded when purchased/acquired.

Form 3 (pages 14-15)

- All payments made by household members for services such as car insurance, telephone bills, DStv, bus and taxi fares, etc. whether paid for when receiving the service or paid monthly.

  Households to record the amount paid for the service.

SEE ALSO THE EXAMPLES ON THE NEXT TWO PAGES

NOT TO BE RECORDED IN THE DIARY:

- Items purchased for business purposes.

HOW TO COMPLETE THE DIARY
Instructions to the main respondent:

- The household member who knows the most about the household's acquisitions should take responsibility for completing the Household Diary.
- A notebook should be issued to individual household members to be carried when away from home, in order to record acquisitions which take place during the day. After completion of a week, tear out the used pages from the notebook and put inside the envelope provided, seal the envelope and give it to the interviewer at his/her next visit.
- Ask each member of the household about any transactions for the day which have not been recorded in the Household Diary.
- Please use the checklist as a reminder regarding items which are easily forgotten.
- ***In order to ensure complete recording, please keep receipts from all purchases.***

FORM 1

ITEMS *PURCHASED* BY THE HOUSEHOLD, OR *RECEIVED* AS GIFTS OR MAINTENANCE, OR *CONSUMED*

Note: *Items purchased from restaurants, fastfood outlets, etc. should be recorded in Form 2*
PAYMENTS FOR *SERVICES* (e.g. *taxi fares, airtime, telephone bills, internet and DStv subscriptions,*

FOR OFFICE USE					Value		
I	COICOP	CHK	Day	Description	Rand	Cent	
			0 1	Brown bread	8	5 0	101
			0 1	Long life full cream milk	7	0 0	102
			0 1	Newspaper	4	5 0	103
			0 1	Milk chocolate	1 1	9 5	104
			0 2	Rice	1 7	9 5	105
			0 5	Women's skirt	2 9 9	9 5	106
			0 5	Fresh potatoes	7	0 0	107
			0 6	Fish caught by household	2 0	0 0	108

	Source *1 = Shop* *2 = Own business* *3 = Own produce (from own garden, livestock, etc.)* *4 = Gift or maintenance received or from nature (hunted, gathered, fished, etc.)* *5 = Other*	**Was this for this household's own consumption?** *1 = Yes* *2 = No, gift given away* *3 = No, maintenance given away*	**Area of purchase** *1 = In a big city (Metro)* *2 = In another urban area (town/city/township)* *3 = In a rural area* 4 = NOT APPLICABLE 5 = DON'T KNOW	**Type of retailer** **Formal sector**: *1 = Chain store*, *2 = Internet*, *3 = Other retailer* **Informal sector**: *4 = Street trading*, *5 = Other* 6 = NOT APPLICABLE 7 = DON'T KNOW	***FOR OFFICE USE*** S	D
101	1	1	1	1		
102	1	1	1	1		
103	1	1	1	3		
104	4	1	2	4		
105	1	1	1	1		
106	1	2	4	6		
107	5	1	2	4		
108	5	1	4	6		

A.2 Assignment of Places of Purchase to Formality Status

This study involves the assignment of a formality status to the places of purchase (PoP) detailed in each line-item of the household consumption diaries. We structured the formality status classification around a number of guiding principles which we detail in this appendix. Although the number and names of the PoP can vary from one country to another, we aimed to limit the number of country-specific choices, and we provide in this appendix a transparent explanation for our choices. We drew upon two sources of information to guide our choices (i) for selected countries we used data from firm censuses to establish a relation between formality and store types or store sizes, and (ii) the International Price Comparison (ICP) project, which builds purchasing power parity indexes and provides a store type classifier for marketed consumption.

At a high level, we use the place of purchase to break down consumption into four categories:

(1) **Non-market consumption** ⇒ Assigned to *informal*

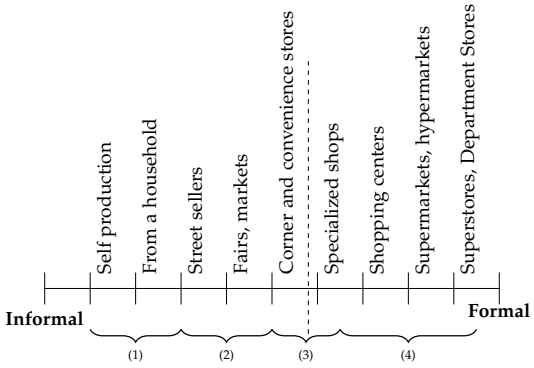
(2) **Market consumption, non brick-and-mortar stores** ⇒ Assigned to *informal*

(3) **Market consumption, small brick-and-mortar stores** ⇒ Depends on scenario. In core scenario (A) this category is broken down: corner stores are assigned to *informal* while specialized stores are assigned to *formal*.

(4) **Market consumption, large brick-and-mortar stores** ⇒ Assigned to *formal*

Based on the above classification, Figure A.2 displays the most common place of purchase names, with our "formality" ordering.

Figure A.2: High-level Overview of Places of Purchase



Using this general categorization and drawing from the place of purchase definitions developed by the International Price Comparison project,[14] we assign the different places of purchase to categories (1) - (4) above. All PoP classified as (1) Non-market and (2) Market, non brick-and-mortar, are assigned to *informal*. These include self production, transfers between households, street selling, fairs, and markets. The reasoning is that small-scale retailers that operate out of temporary settlements are more likely to be informal and less likely to pay taxes than businesses with permanent, regulated locations. In contrast, all PoP classified as (4) Market consumption from large brick-and-mortar, such as supermarkets, shopping centers, and department stores are assigned to *formal*. These larger institutions, by virtue of their size, number of employees, and establishment size, among other factors, are likely to be registered and pay commodity taxes.

Category (3) - Market consumption from small brick-and-mortar stores - is the category for which the formality status might be most context dependent. This category spans corner stores which tend to be small, but also specialized shops which might be more heterogeneous in size. In our central scenario (A) we assign corner and convenience stores to informal and specialized stores to formal. In Appendix A.2 we report results for two alternative scenarios: Scenario (B) where we assign all of category (3) to informal, and scenario (C), where we assign all of category (3) to formal. Table A4 outlines scenarios (A) to (C), as well as a comparison between our classification and the ICP classification for the most common places of purchases across the 20 countries.

While places of purchases in categories (1)-(4) characterize a large share of expenditure for the countries in our sample, some places of purchase do not fit in this nomenclature and are assigned to category (5) "other places of purchases". We list below the most common occurrences:

(5) **other places of purchase** $\Rightarrow$ Typically assigned to *formal*
We list below some common occurrences:

- public institutions
- health, education, and financial institutions
- abroad
- online
- entertainment (hotels, restaurants, shows)

[14] See Table A4 for details on the International Price Comparison project's classification

As can be appreciated from the list above, most of these PoP are likely to be formal and taxable.[15] We assign these PoPs to formal unless there are additional details in the labels of the PoP about their relative size: for example, within health spending, there may be categories called "public health institution," versus "traditional medicine." In this example, the latter would be considered informal and the former two would be considered as formal. Similarly, within entertainment, we assign formal to "restaurant" and informal to "canteen-truck." We provide further details on these nuanced scenarios in Tables A4 and A5.

[15]Whether these place of purchases are taxed in practice, is a different matter. For example taxing online purchases is not technically difficult, but tax systems have been slow to adjust, in part due to issues of assessing the correct jurisdiction where the tax should be remitted.

Table A4: Store Types: ICP vs. Paper's Classification

Place of Purchase	Our Classification	ICP Classification[a]	Scenario A	Scenario A	Scenario C
Self production	(1)	-	Informal	Informal	Informal
From a household	(1)	-	Informal	Informal	Informal
From the bush, forest	(1)	-	Informal	Informal	Informal
Street sellers	(2)	4	Informal	Informal	Informal
Mobile cart	(2)	4	Informal	Informal	Informal
Fairs, open markets	(2)	3	Informal	Informal	Informal
Informal, irregular transport	(2)	-	Informal	Informal	Informal
Traditional medicine	(2)	-	Informal	Informal	Informal
Individual service provider	(2)	7	Informal	Informal	Informal
Plumbing services	(2)	7	Informal	Informal	Informal
Epiceries, bodegas	(3)	2	Informal	Informal	Formal
Corner shops	(3)	2	Informal	Informal	Formal
Convenience stores	(3)	2	Informal	Informal	Formal
Warehouses	(3)	5	Informal	Informal	Formal
Tourism agencies	(3)	7	Formal	Informal	Formal
Electronic, car repair	(3)	7	Formal	Informal	Formal
Fruit, vegetable store	(3)	6	Formal	Informal	Formal
Rotisserie, Butcher	(3)	6	Formal	Informal	Formal
Bakery, Patisserie	(3)	6	Formal	Informal	Formal
Furniture shop	(3)	6	Formal	Informal	Formal
Electronics, music store	(3)	6	Formal	Informal	Formal
Book stores, photocopy center	(3)	6	Formal	Informal	Formal
Pharmacies	(3)	6	Formal	Informal	Formal
Canteens, fast-food	(5)	-	Informal	Informal	Informal
Supermarkets, hypermarkets	(4)	1	Formal	Formal	Formal
Department stores	(4)	1	Formal	Formal	Formal
Hotels	(5)	7	Formal	Formal	Formal
Restaurants, bars	(5)	7	Formal	Formal	Formal
Public administration	(5)	8	Formal	Formal	Formal
Public and semi-public agencies	(5)	8	Formal	Formal	Formal
Public education institution	(5)	8	Formal	Formal	Formal
Public medical institution	(5)	8	Formal	Formal	Formal
Private medical institution	(5)	7	Formal	Formal	Formal
Private education institution	(5)	7	Formal	Formal	Formal
Banks, financing institutions, bars	(5)	-	Formal	Formal	Formal
Professional legal services	(5)	7	Formal	Formal	Formal
Abroad	(5)	-	Formal	Formal	Foraml
Online	(5)	9	Formal	Formal	Formal

[a]Outlet types, based on the ICP classification are **1** Large shops (Hypermarkets, Supermarkets, Department stores); **2** Medium and small shops (Minimarkets, Kiosks, Neighborhood shops, Grocery stores, Convenience stores); **3** Markets (open markets, covered markets, wet markets); **4** Street outlets (Mobile shops, Street vendors); **5** Bulk wholesale stores and discount shops; **6** Specialized shops (Supply shops, hardware shops, furniture stores); **7** Private service providers (taxi cabs, hotels, restaurants, private schools, private hospitals); **8** Public or semi-public service providers (water suppliers, electric power companies, public schools, hospitals); and **9** Other kinds of trade (online, Internet shopping, catalogue orders).

Table A5: Country-Specific Places of Purchase

	BRAZIL				BURKINA FASO			
Sector			new classification	recoded as			new classification	recoded as
Formal	2.70%	restaurant	5: other consumption	entertainment	2.10%	Autres service prives	5: other consumption	institutions
	0.70%	bank	5: other consumption	institutions	1.40%	Service de transport prive	5: other consumption	institutions
	2.10%	bar-caf_	5: other consumption	entertainment	1.80%	Bar, cafe, restaurant, hotel	5: other consumption	entertainment
	1.50%	health institution	5: other consumption	institutions	0.65%	Ecole, lycee, universite publics	5: other consumption	public sector
	2.50%	education institution	5: other consumption	institutions	0.61%	Cabine telephone privee	5: other consumption	institutions
	0.60%	internet	5: other consumption	internet	1.00%	Clinique, laboratoire medical public	5: other consumption	public sector
	11.50%	supermarket	4: market, larger stores	large stores	1.98%	Telephone, eau, electricite	5: other consumption	public sector
	0.70%	department store	4: market, larger stores	large stores	1.17%	Ecole, lycees, universite privas	5: other consumption	institutions
	3.40%	grocery store	3: market, small store-front	specialized stores	0.57%	Magasin de gros a petits prix	4: market, larger stores	large stores
	22.10%	specialized shop	3: market, small store-front	specialized stores	0.69%	Grands magasin	4: market, larger stores	large stores
	4.00%	pharmacy	3: market, small store-front	specialized stores	1.69%	Station service (lubrifiants)	3: market, small store-front	specialized stores
	6.50%	vehicle	3: market, small store-front	specialized stores	1.30%	Atelier, service reparation	3: market, small store-front	specialized stores
					1.03%	Pharmacie	3: market, small store-front	specialized stores
					0.53%	Quincallerie (petite taille)	3: market, small store-front	specialized stores
Informal	8.30%	person	2: market, no store front	individual	1.43%	Boutique de quartier	3: market, small store-front	convenience stores
	1.40%	street seller	2: market, no store front	street selling	0.88%	Kiosque ou echoppe quartier	2: market, no store front	street selling
	1.20%	fair	2: market, no store front	street selling	0.98%	Marchant ambulants	2: market, no store front	street selling
	1.40%	small market	2: market, no store front	street selling	4.05%	Marche	2: market, no store front	street selling
	4.80%	private service	2: market, no store front	individual	1.24%	Menage	1: non-market	from a household/transfers
	0.80%	small shop	2: market, no store front	street selling	1.01%	Bien ou service autoproduit	1: non-market	self production
					0.51%	Cadeau recu en nature ou en espace	1: non-market	from a household/transfers
Unspecified	20.80%	unspecified		not applicable/other				

	BURUNDI				CAMEROON			
Sector			new classification	recoded as			new classification	recoded as
Formal	2.10%	Secteur public ou parapublic	5: other consumption	public sector	7.00%	Hotels/bars/restaurants	5: other consumption	entertainment
	3.60%	Autre lieu d'achat formel	4: market, larger stores	large stores	2.90%	Presetation de services publics	5: other consumption	public sector
	0.80%	Magasin, atelier formel (societe)	3: market, small store-front	specialized stores	2.10%	Cliniques	5: other consumption	institutions
					7.50%	Secteur transport	5: other consumption	public sector
					1.00%	Supermarche/Grand magasin	4: market, larger stores	large stores
					3.80%	Magasin specialistes	3: market, small store-front	specialized stores
Informal	3.87%	Vendeur ambulant	2: market, no store front	street selling	2.50%	Prestation de services individuels	5: other consumption	individual
	13.87%	Autre lieu d'achat informel	2: market, no store front	other informal	10.70%	Epiceries/Boutiques/Echoppes	3: market, small store-front	corner shops
	32.92%	Marche public	2: market, no store front	street selling	0.80%	Vendeurs specialises hors magasins	2: market, no store front	street selling
	26.84%	Domicile du vendeur	1: non-market	other informal	26.40%	Marches	2: market, no store front	street selling
	14.30%	Bien ou service autoproduit	1: non-market	self production	3.40%	Kiosque de jeux et Call Box	2: market, no store front	street selling
	0.86%	Cadeau Recu	1: non-market	from a household/transfers	3.40%	Vente ambulante	2: market, no store front	street selling
					1.90%	Domicile de vendeur	1: non-market	from a household/transfers
					0.90%	Dans la nature/forit/brousse	1: non-market	self production
					3.60%	Auto production	1: non-market	self production
					14.80%	Don, cadeau recu	1: non-market	from a household/transfers
Unspecified					7.30%	Autre		not applicable/other

	CHILE				COLOMBIA		
Sector			new classification	recoded as		new classification	recoded as
Formal	1.98%	INTERNET	5: other consumption	internet	5.08% Restaurantes	5: other consumption	entertainment
	2.37%	HOSPITAL PñBLICO Y CONSULTORIOS	5: other consumption	public sector	1.05% Cafeteras y establecimientos de comidas rpidas	5: other consumption	entertainment
	5.39%	CLeNICAS	5: other consumption	public sector	0.96% Televentas y ventas por catlogo	5: other consumption	internet
	0.96%	RESTAURANTES Y BARES	5: other consumption	entertainment	0.00% A travs de Internet	5: other consumption	internet
	4.29%	DISTRIBUIDORAS - MAYORISTAS	4: market, larger stores	large stores	10.07% Almacenes o supermercados de cadena y tiendas por departamento	4: market, larger stores	large stores
	26.55%	SUPERMERCADOS	4: market, larger stores	large stores	2.05% Plazas de mercado y galeras	4: market, larger stores	large stores
	4.85%	FARMACIAS	3: market, small store-front	specialized stores	0.55% Hipermercados	4: market, larger stores	large stores
	0.53%	TIENDA ESPECIALIZADA	3: market, small store-front	specialized stores	11.12% Establecimientos especializados en la venta del artculo o la prestacin del servicio adquirido	3: market, small store-front	specialized stores
	1.61%	FERRETEReAS Y MULTIFERRETEReAS	3: market, small store-front	specialized stores	1.73% Farmacias y drogueras	3: market, small store-front	specialized stores
Informal	13.30%	ALMACN TRADICIONAL	3: market, small store-front	corner shops	0.85% Graneros	3: market, small store-front	corner shops
	2.85%	COMERCIO AMBULANTE	2: market, no store front	street selling	4.72% Supermercados de barrio	3: market, small store-front	corner shops
	3.27%	FERIAS LIBRES	2: market, no store front	street selling	13.55% Tiendas de barrio	3: market, small store-front	corner shops
	0.93%	VEGAS - MERCADOS	2: market, no store front	street selling	1.13% Persona particular	2: market, no store front	individual
					1.71% Vendedores ambulantes o ventas callejeras	2: market, no store front	street selling
					0.89% Transfers, from household	1: non-market	from a household/transfers
					10.31% Self production	1: non-market	self production
Unspecified	29.33%	NA		not applicable/other	10.58% [Unspecified]		not applicable/other
	0.93%	ESTABLECIMIENTOS IMPOSIBLES DE IDENTI..		not applicable/other	21.98% [Missing]		missing

	COSTA RICA				DEM. REPUBLIC of CONGO			
Sector			new classification	recoded as			new classification	recoded as
Formal	3.86%	Restaurante / soda / cafetera / heladera	5: other consumption	entertainment	3.07%	Achat secteur public	5: other consumption	public sector
	1.71%	Comedor en lugar de trabajo	5: other consumption	entertainment	0.54%	Achat supermarche	4: market, larger stores	large stores
	1.06%	En el exterior	5: other consumption	abroad	3.19%	Achat magasin indo-pakistanais	3: market, small store-front	specialized stores
	1.86%	Laboratorio / clnica / centro medico	5: other consumption	institutions	3.77%	Achat magasin non indo-pakistanais	3: market, small store-front	specialized stores
	17.11%	Supermercado	4: market, larger stores	large stores				
	1.22%	Tienda por departamentos	4: market, larger stores	large stores				
	3.42%	Tienda de ropa / zapatera / perfumera	3: market, small store-front	specialized stores				
	4.26%	Gasolinera y estacion de servicio	3: market, small store-front	specialized stores				
	1.10%	Carnicera / pescadera	3: market, small store-front	specialized stores				
	3.39%	Almacn de electrodomosticos y de tecnologas	3: market, small store-front	specialized stores				
	11.34%	Local especializado	3: market, small store-front	specialized stores				
	0.98%	Salones de esttica o belleza	3: market, small store-front	specialized stores				
Informal	6.21%	Pulperia o minisuper	3: market, small store-front	corner shops	10.07%	Achat Ambulant	2: market, no store front	street selling
	9.06%	Recibido o comprado a otros hogares	2: market, no store front	street selling	36.48%	Achat marche public	2: market, no store front	street selling
	0.80%	Local de artculos usados	2: market, no store front	street selling	5.76%	Autre lieu informel	2: market, no store front	other informal
	2.41%	Vendedor ambulante o a domicilio	1: non-market	self production	17.88%	Achat domicile	1: non-market	self production
	0.80%	Retiro del negocio	1: non-market	self production	17.53%	Bien ou service autoproduit	1: non-market	self production
					1.38%	Cadeau recu	1: non-market	from a household/transfers
Unspecified	3.72%	Otro		not applicable/other				
	1.38%	Imputado		not applicable/other				
	20.03%	[Missing]		missing				

	ECUADOR				MEXICO			
Sector			new classification	recoded as			new classification	recoded as
Formal	6.61%	Restaurantes, salones	5: other consumption	entertainment	0.72%	Diconsa	5: other consumption	public sector
	2.61%	Transporte de pasajeros	5: other consumption	public sector	2.62%	Loncherias, fondas, torterias , cocinas economicas, cenadurias	5: other consumption	entertainment
	0.74%	Establecimientos educativos	5: other consumption	institutions	2.35%	Restaurantes	5: other consumption	entertainment
	0.74%	Hipermercados	4: market, larger stores	large stores	2.08%	Tiendas departamentales	4: market, larger stores	large stores
	2.10%	Supermercados de cadena	4: market, larger stores	large stores	11.37%	Supermercados	4: market, larger stores	large stores
	0.63%	Ropa de todo tipo	3: market, small store-front	specialized stores	0.50%	Compras fuera del pai_s	4: market, larger stores	large stores
	1.08%	Tercena/carnicera	3: market, small store-front	specialized stores	0.97%	Tiendas con membresi_a	4: market, larger stores	large stores
	0.89%	Boticas y farmacias	3: market, small store-front	specialized stores	21.11%	Tiendas especi_ficas del ramo	3: market, small store-front	specialized stores
	5.16%	Panaderas	3: market, small store-front	specialized stores				
	2.35%	Gasolineras	3: market, small store-front	specialized stores				
	0.64%	Personas particulares	2: market, no store front	individual				
Informal	0.96%	Bodegas, distribuidores	3: market, small store-front	corner shops	0.58%	Tiendas de conveniencia	3: market, small store-front	corner shops
	30.38%	Tiendas de barrio	3: market, small store-front	corner shops	12.81%	Tiendas de abarrotes	3: market, small store-front	corner shops
	12.89%	Mercados	2: market, no store front	street selling	3.11%	Vendedores ambulantes	2: market, no store front	street selling
	0.62%	Ferias libres	2: market, no store front	street selling	5.61%	Persona particular	2: market, no store front	individual
	2.56%	Vendedores ambulantes	2: market, no store front	street selling	3.65%	Mercado	2: market, no store front	street selling
	23.04%	Productos autoconsumo, autosuministro	1: non-market	self production	1.99%	Tianguis o mercado sobre ruedas	2: market, no store front	street selling
					1.33%	Auto produccin	1: non-market	self production
Unspecified	1.83%	Otros		not applicable/other	28.79%	No aplica		not applicable/other

	MOROCCO				MOZAMBIQUE			
Sector			new classification	recoded as			new classification	recoded as
Formal	35.48%	Public and semi-public agencies	5: other consumption	public sector	8.80%	loja	3: market, small store-front	specialized stores
	4.11%	Private education institution	5: other consumption	institutions				
	1.50%	Regular transportation means (bus, train, plane)	5: other consumption	public sector				
	0.66%	Public baths, shower, swimming pool	5: other consumption	public sector				
	1.30%	Medical care in private institution	5: other consumption	institutions				
	1.69%	Public administration	5: other consumption	public sector				
	0.84%	Modern clothing shop	3: market, small store-front	specialized stores				
	1.83%	Pharmacy	3: market, small store-front	specialized stores				
	0.53%	Small Bookshop, kiosk	3: market, small store-front	specialized stores				
	1.22%	Craftsman's shop (hairdresser, tailor)	3: market, small store-front	specialized stores				
	3.12%	Butcher or retail chicken seller	3: market, small store-front	specialized stores				
	0.65%	Gas stations (benzine)	3: market, small store-front	specialized stores				
Informal	0.73%	Cafe, non-standing restaurant	5: other consumption	informal entertainment	18.65%	mercado informal	2: market, no store front	street selling
	9.69%	Neighborhood or village grocer	3: market, small store-front	corner shops	12.24%	mercado	2: market, no store front	street selling
	1.85%	Grocer's	3: market, small store-front	corner shops	31.49%	auto produo	1: non-market	self production
	2.63%	Neighborhood market	2: market, no store front	street selling				
	0.52%	City market or central market	2: market, no store front	street selling				
	10.73%	Weekly market	2: market, no store front	street selling				
	1.34%	Itinerant merchant selling on sidewalks	2: market, no store front	street selling				
	3.62%	Self production/consumption	1: non-market	self production				
Unspecified	4.70%	Other places		not applicable/other	9.65%	missing		missing
	0.66%	Unknown		not applicable/other	18.83%	outro		not applicable/other
	5.81%	Not relevant		not applicable/other				

	NIGER				PAPUA NEW GUINEA			
Sector			new classification	recoded as			new classification	recoded as
Formal	7.27%	Prestation services publiques	5: other consumption	public sector	34.45%	Supermarket	4: market, larger stores	large stores
	0.84%	Secteur transport	5: other consumption	public sector				
	0.91%	Hotel, bar restaurant	5: other consumption	entertainment				
Informal	45.54%	Epicerie, boutique	3: market, small store-front	convenience store	9.35%	Small shop, canteen, tuck shop	3: market, small store-front	corner shops
	4.81%	Vente ambulante	2: market, no store front	street selling	10.52%	Local market	2: market, no store front	street selling
	20.19%	Marche	2: market, no store front	street selling	3.76%	Street vendor	2: market, no store front	street selling
	4.93%	Prestation service individuels	2: market, no store front	individual	10.17%	Gift	1: non-market	from a household/transfers
	4.41%	Auto production	1: non-market	self production	14.17%	Home production	1: non-market	self production
	0.41%	Cadeau recu	1: non-market	from a household/transfers				
Unspecified	6.61%	Autre		not applicable/other	17.57%	Other		not applicable/other

	PERU				REPUBLIC of CONGO			
Sector			new classification	recoded as			new classification	recoded as
Formal	3.42%	Empresas de Transporte formales	5: other consumption	public sector	3.94%	Hotels, restaurants, bars, cafes	5: other consumption	entertainment
	1%	Talleres formales	5: other consumption	institutions	2.51%	Cliniques, laboratoires mdicaux et coles	5: other consumption	institutions
	0.62%	Clnica particular	5: other consumption	institutions	3.94%	Secteur transports	5: other consumption	public sector
	1.56%	Centro de estudios	5: other consumption	institutions	5.84%	Prestataires de services publics	5: other consumption	public sector
	0.96%	Restaurantes y/ bares	5: other consumption	entertainment	1.03%	Grands magasins	4: market, larger stores	large stores
	1.33%	Grifos de empresas	5: other consumption	public sector	3.43%	Epiceries modernes	3: market, small store-front	specialized stores
	1.05%	Bodega (por mayor)	4: market, larger stores	large stores	6.98%	Autres commerces modernes	3: market, small store-front	specialized stores
	3.42%	Supermercado	4: market, larger stores	large stores				
	3.63%	Farmacia	3: market, small store-front	specialized stores				
	5.78%	Tienda especializada al por menor	3: market, small store-front	specialized stores				
	9.12%	Librera	3: market, small store-front	specialized stores				
	0.79%	Panadera	3: market, small store-front	specialized stores				
	0.65%	Peluquera	3: market, small store-front	specialized stores				
Informal	14.62%	Bodega (por menor)	3: market, small store-front	corner shops	42.78%	Marches	2: market, no store front	street selling
	23.98%	Mercado (por menor)	2: market, no store front	street selling	6.17%	Marchands ambulants	2: market, no store front	street selling
	2.67%	Feria	2: market, no store front	street selling	8.44%	Echoppes sur marches et sur bord de route	2: market, no store front	street selling
	4.97%	Ambulante	2: market, no store front	street selling	5.50%	Prestataires de services individuels	2: market, no store front	individual
	3.34%	Mercado (por mayor)	2: market, no store front	street selling	4.55%	Produit autoconsommes	1: non-market	self production
					3.93%	Menages	1: non-market	self production
Unspecified	22.31%	Otro (Especifique)		not applicable/other				

	RWANDA				SOUTH AFRICA			
Sector			new classification	recoded as			new classification	recoded as
Formal	2.44%	Bar/restaurant	5: other consumption	entertainment	38.64%	Chain store	4: market, larger stores	chain stores
	0.59%	Supermarket/big shop	4: market, larger stores	large stores	11.18%	Other retailer	3: market, small store-front	other retailers
	4.63%	Specialized shop	3: market, small store-front	specialized stores				
Informal	13.55%	Small shop/boutique	3: market, small store-front	corner shops	2.72%	Other	2: market, no store front	other informal
	13.14%	Service provider	2: market, no store front	individual	0.88%	Street trading	2: market, no store front	street selling
	0.77%	Mobile seller	2: market, no store front	street selling	0.63%	From a household	1: non-market	from a household/transfers
	1.72%	Individual	2: market, no store front	individual				
	12.47%	Market	2: market, no store front	street selling				
	11.49%	Self production	1: non-market	self production				
	26.49%	From a household	1: non-market	from a household/transfers				
Unspecified	12.71%	Other		not applicable/other	0.55%	Not applicable		not applicable/other
					45.11%	Unspecified		not applicable/other

	TANZANIA				URUGUAY			
Sector			new classification	recoded as			new classification	recoded as
Formal	37.49%	Shop	3: market, small store-front	specialized stores	0.77%	Bar, Pizzeria	5: other consumption	entertainment
					0.52%	Cantina Trabajo Colegio	5: other consumption	entertainment
					0.85%	Restaurante, Parrillada	5: other consumption	entertainment
					11.73%	Autoservicio, Cadena de Supermercados	4: market, larger stores	large stores
					0.76%	Fuera del pais	4: market, larger stores	large stores
					0.97%	Shopping o galeria	4: market, larger stores	large stores
					2.29%	Merceria, Tienda	3: market, small store-front	specialized stores
					2.58%	Carniceria, Polleria, Pescaderia	3: market, small store-front	specialized stores
					1.49%	Panaderia, Confiteria	3: market, small store-front	specialized stores
					0.66%	Verduleria, Puesto, Fruteria	3: market, small store-front	specialized stores
					0.90%	Zapateria, Marroquineria, Talabarteria	3: market, small store-front	specialized stores
					1.28%	Casa de electrodomesticos, telefonos	3: market, small store-front	specialized stores
					0.71%	Farmacia, Perfumeria, Panalera	3: market, small store-front	specialized stores
Informal	2.44%	Street vendor	2: market, no store front	street selling	0.75%	Quiosco, Salon	3: market, small store-front	corner shops
	22.64%	Market	2: market, no store front	street selling	7.76%	Almacen	3: market, small store-front	corner shops
	4.73%	Other household	1: non-market	from a household/transfers	1.50%	Feria vecinal	2: market, no store front	street selling
	1.82%	Gift or free	1: non-market	from a household/transfers	1.01%	Vendedor ambulante, Puesto callejero, Carrito	2: market, no store front	street selling
	15.90%	Produced by household	1: non-market	self production				
Unspecified	13.77%	Other		not applicable/other	11.66%	Missing		missing
					47.97%	No corresponde		not applicable/other

B Theory Appendix

B.1 Allowing for cross-price elasticities across commodities

Writing ϵ_{kj} the (uncompensated) cross-price elasticity of demand across commodities j and k we can write the optimal tax rates in the absence of an informal sector τ_j^* in the following way:

$$\tau_j^* = \frac{(\mu - g) - \int_i (g^i - g)\beta^i s_j^i / s_j di + \mu \sum_{k \neq j} \tau_k \epsilon_{k,j} s_k / s_j}{-\mu \epsilon_j} \tag{10}$$

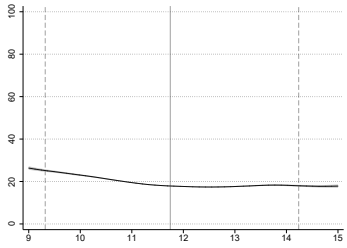
Optimal tax rates when the informal sector cannot be taxes can similarly be written as

$$\tau_j^{**} = \frac{(\mu - g) - \int_i (g^i - g)\beta^i s_{j1}^i / s_{j1} + \mu \sum_{k \neq j} \tau_k \epsilon_{k1,j1} s_{k1} / s_{j1}}{-\mu \epsilon_{j1}} \tag{11}$$

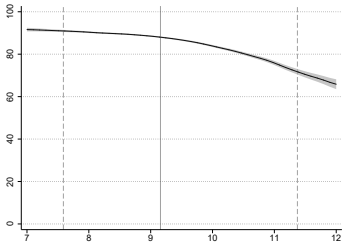
The terms $\mu \sum_{k \neq j} \tau_k \epsilon_{k,j} s_k / s_j - \mu \epsilon_j$ and $\sum_{k \neq j} \tau_k \epsilon_{k1,j1} s_{k1} / s_{j1} - \mu \epsilon_{j1}$ indicate that commodities that are close substitutes should be taxed at similar rates. All the intuitions discussed above regarding the differences between τ_j^* and τ_j^{**} still hold when cross-price elasticities across commodities are non-zero.

C Additional Figures and Tables

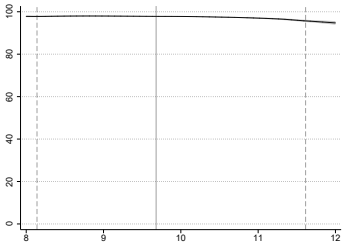
Figure C.1: Informality Engel Curves



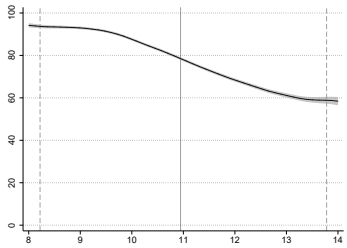
(a) Brazil



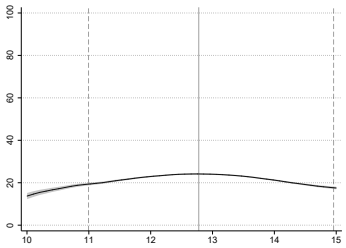
(b) Burkina Faso



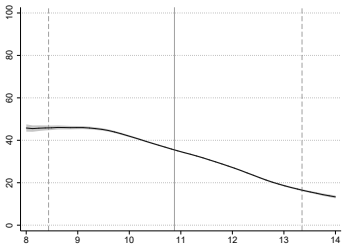
(c) Burundi



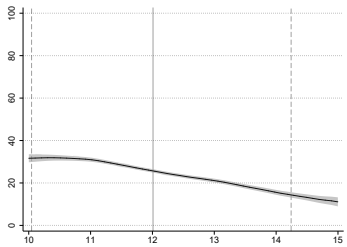
(d) Cameroon



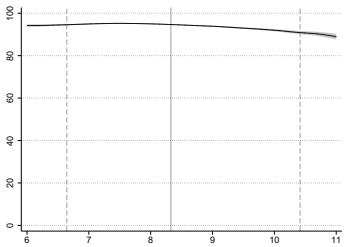
(e) Chile



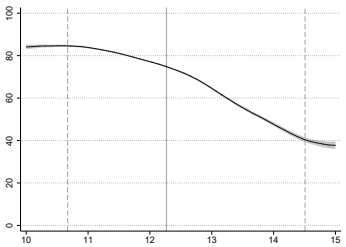
(f) Colombia



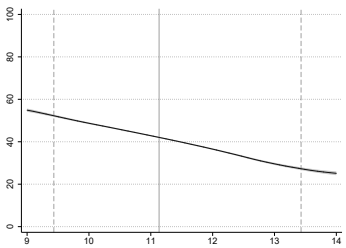
(g) Costa Rica



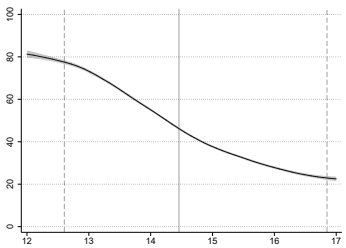
(h) Dem. Rep. of Congo



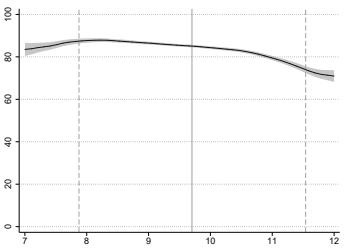
(i) Ecuador



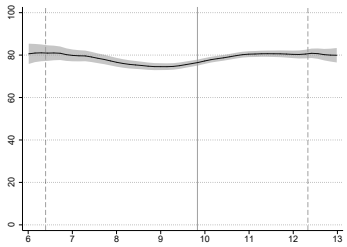
(j) Mexico



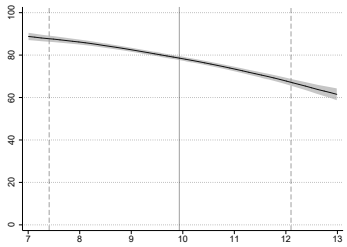
(k) Morocco



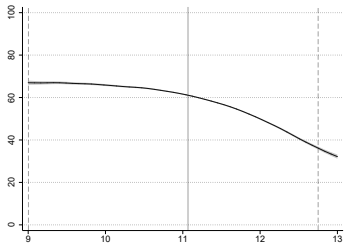
(l) Mozambique



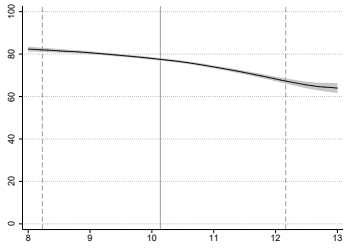
(m) Niger



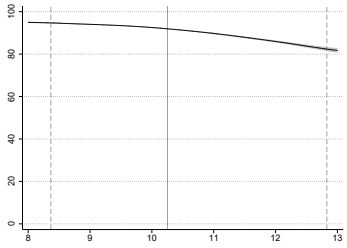
(n) Papua New Guinea



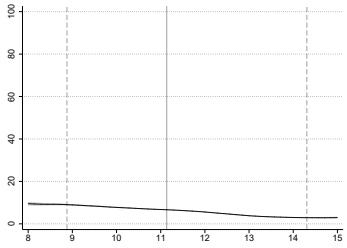
(o) Peru



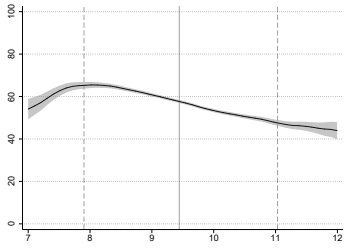
(p) Rep. of Congo



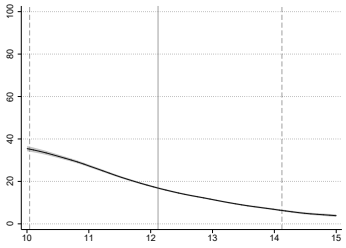
(q) Rwanda



(r) South Africa



(s) Tanzania



(t) Uruguay

Local polynomial fit of the informality Engel curves for DR Congo and Mexico as presented in section 3.1. The informality Engel curves are defined as the share of informal consumption over the per capita expenditure of households. Expenditure is measured in log base 2, such that a one unit increase on the horizontal axis corresponds to a doubling of household's expenditure. The shaded area around the polynomial fit corresponds to the 95% confidence interval. The solid grey line corresponds to the median of each country's expenditure distribution, while the dotted lines correspond to the 5th and 95th percentiles.

Figure C.2: COICOP Product Composition[16]

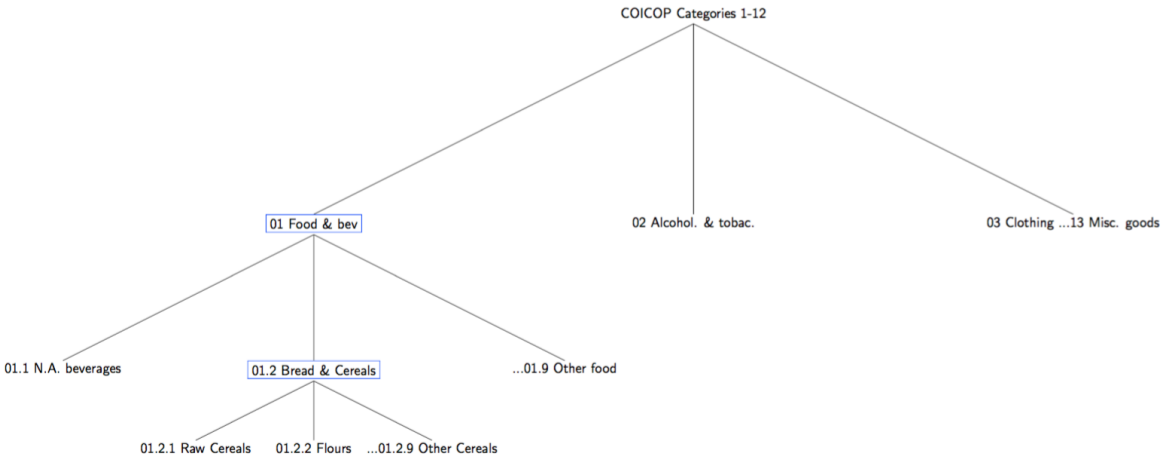
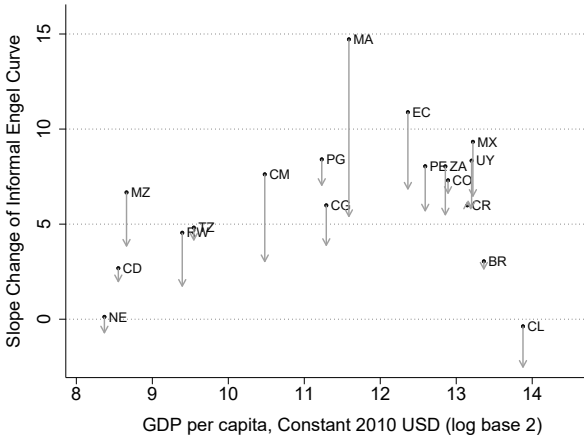
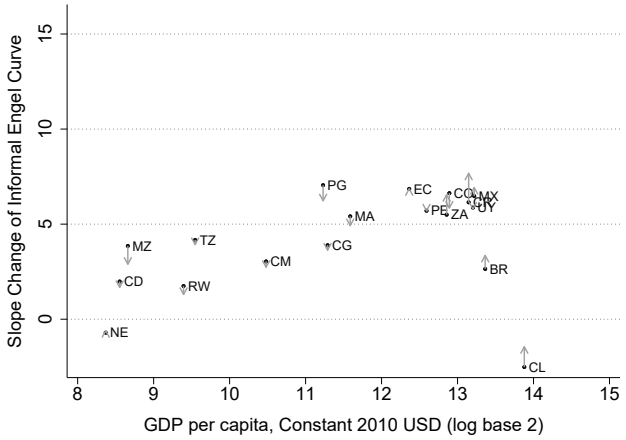


Figure C.3: Change in IEC Slopes from Product Composition

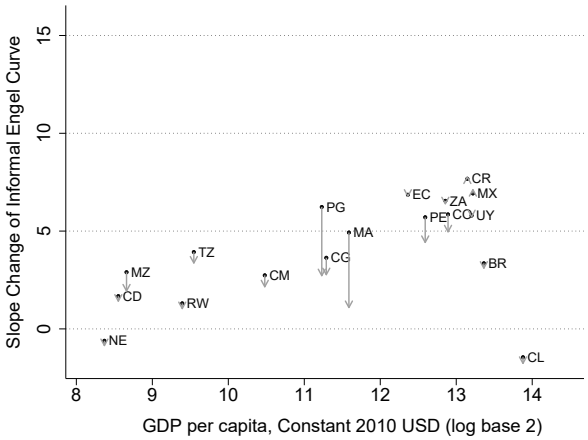
(a) No product controls to COICOP 2



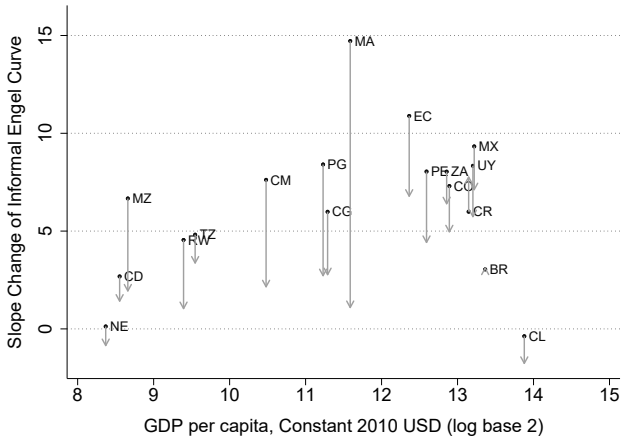
(b) COICOP 2 to COICOP 3



(c) COICOP 3 to COICOP 4

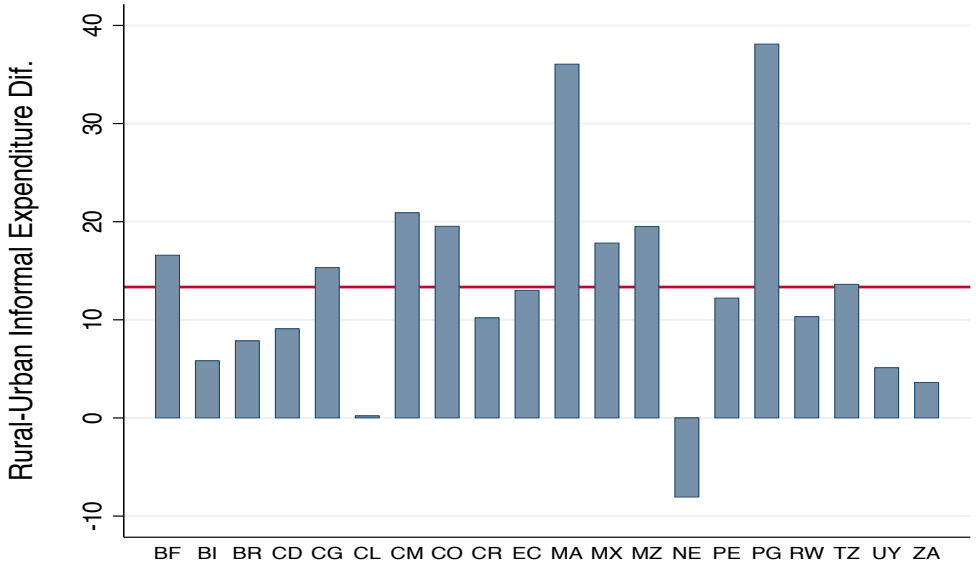


(d) No product controls to COICOP 4



This figure shows the change in the informal Engel curves' slopes when controlling for increasingly narrow product code levels.

Figure C.4: Rural-Urban difference in Informal Expenditure Shares



The figure shows the difference between the rural and urban share of informal expenditure across countries. The red line shows the sample average of 13%.

Table C1: Main Regression Results

	Main		Geography		Product Codes			District + Products		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	No control	Hhld Char.	Rur/Urb.	District	Level2	Level3	Level4	Level2	Level3	Level4
Brazil	1.92	2.02	1.73	1.38	1.38	3.05	3.05	.8	2.59	2.64
	(.06)	(.06)	(.06)	(.06)	(.06)	(.05)	(.05)	(.06)	(.06)	(.05)
Burkina Faso	5.7	6.01	4.03	5.47	3.68	3.47	2.41	3.52	3.3	2.34
	(.18)	(.19)	(.18)	(.19)	(.12)	(.11)	(.11)	(.12)	(.11)	(.11)
Burundi	1.69	2.05	1.48	1.19	1.31	.99	.72	.96	.68	.49
	(.12)	(.13)	(.12)	(.12)	(.07)	(.05)	(.04)	(.07)	(.06)	(.05)
Cameroon	6.85	7.13	5.99	5.81	3.17	2.98	2.52	2.51	2.44	2.01
	(.12)	(.13)	(.17)	(.14)	(.11)	(.1)	(.1)	(.12)	(.11)	(.1)
Chile	.61	.55	.55	.2	-1.86	-1.98	-2.54	-1.92	-1.95	-2.37
	(.09)	(.1)	(.1)	(.11)	(.08)	(.07)	(.06)	(.09)	(.08)	(.06)
Colombia	7.11	6.75	6.36	5.94	6.6	6.21	5.44	6.26	5.89	5.15
	(.09)	(.1)	(.1)	(.1)	(.07)	(.07)	(.06)	(.07)	(.07)	(.06)
Costa Rica	5.04	5.99	5.33	3.93	6.2	7.79	7.88	4.57	6.38	6.46
	(.22)	(.24)	(.24)	(.28)	(.2)	(.19)	(.15)	(.24)	(.22)	(.18)
Dem. Rep. Congo	1.64	2.85	2.07	2.82	2.09	1.76	1.49	2.1	1.76	1.52
	(.11)	(.12)	(.12)	(.13)	(.09)	(.08)	(.06)	(.1)	(.09)	(.07)
Ecuador	10.82	10.8	10.44	10.08	6.85	6.84	6.77	5.98	5.99	6.13
	(.1)	(.11)	(.11)	(.12)	(.09)	(.09)	(.08)	(.1)	(.1)	(.09)
Mexico	9.01	9.9	8.29	6.87	5.83	6.34	6.52	3.14	3.69	4.11
	(.12)	(.14)	(.14)	(.15)	(.12)	(.11)	(.09)	(.13)	(.12)	(.1)
Morocco	14.34	15.27	10.4	14.17	6.99	6.19	2.47	6.31	5.53	2.03
	(.11)	(.12)	(.14)	(.13)	(.1)	(.08)	(.06)	(.1)	(.09)	(.07)
Mozambique	6.25	7.26	5.52	4.98	3.99	3.02	2.09	2.93	2.22	1.56
	(.25)	(.27)	(.27)	(.27)	(.2)	(.19)	(.16)	(.2)	(.19)	(.16)
Niger	.07	-.13	.25	-.32	-.74	-.67	-.83	-.89	-.77	-.91
	(.2)	(.21)	(.25)	(.24)	(.19)	(.18)	(.16)	(.22)	(.21)	(.18)
Papua New Guinea	8.18	8.77	5.4	5.31	7.12	6.49	4.05	4.05	3.63	2.61
	(.33)	(.33)	(.32)	(.33)	(.27)	(.26)	(.2)	(.27)	(.26)	(.21)
Peru	7.22	7.59	7.02	7.16	5.4	5.4	4.25	4.62	4.62	3.72
	(.1)	(.1)	(.12)	(.13)	(.06)	(.06)	(.05)	(.08)	(.08)	(.06)
Rep. of Congo	4.42	5.25	4.65	3.65	3.42	3.31	2.51	2.75	2.65	2.07
	(.21)	(.22)	(.22)	(.23)	(.15)	(.13)	(.11)	(.16)	(.14)	(.12)
Rwanda	3.55	3.53	2.78	3.11	1.2	.98	.93	.94	.66	.63
	(.11)	(.11)	(.11)	(.12)	(.06)	(.05)	(.05)	(.06)	(.05)	(.05)
South Africa	1.39	1.84	1.6	1.69	1.55	1.83	1.84	1.44	1.79	1.81
	(.05)	(.06)	(.06)	(.07)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)
Tanzania	4.8	3.94	3.71	.23	3.55	3.29	2.86	.72	.33	.31
	(.28)	(.29)	(.29)	(.27)	(.24)	(.230	(.19)	(.24)	(.23)	(.2)
Uruguay	8.18	8.47	8.22	7.44	5.12	4.91	4.89	4.34	4.13	4.04
	(.17)	(.19)	(.19)	(.22)	(.23)	(.21)	(.16)	(.22)	(.19)	(.14)
All Countries (Mean)	5.44	5.79	4.79	4.56	3.64	3.61	2.97	2.76	2.78	2.32
Household Characteristics		X	X	X	X	X	X	X	X	X
Urban/Rural			X							
Minimum Geoloc				X				X	X	X
COICOP 2-dig					X			X		
COICOP 3-dig						X			X	
COICOP 4-dig							X			X

This table shows the average slope of the Informal Engel curve across countries for different specifications. The slopes are estimated from: $Share\,Informal_i = \beta.ln(expenditure\,pc)_i + \Gamma X_i + \varepsilon_i$, where the dependent variable is the informal expenditure share and the explanatory variable is the log expenditure pp. Controls include household characteristics (household size, age, gender, and education of head), geographic indicators (urban/rural and districts), and product codes at the 2nd, 3rd and 4th level of the United Nation's COICOP classification. While all countries follow the COICOP at the 2nd level (12 categories: food, clothing etc.), Brazil, Morocco and Peru have specific product classifications at lower levels. The geographic control "district" refers to the lowest geographical level available in each survey.

Table C2: Regression Results - Robustness Scenario B

	Main		Geography		Product Codes			District + Products		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	No control	Hhld Char.	Rur/Urb.	District	Level2	Level3	Level4	Level2	Level3	Level4
Brazil	2.13	2.03	1.62	1.23	0.78	3.46	2.79	0.00	2.92	2.33
	0.07	0.07	0.07	0.07	0.07	0.05	0.04	0.07	0.05	0.04
Burkina Faso	4.08	4.29	2.89	3.92	2.60	2.38	1.67	2.55	2.31	1.68
	0.15	0.16	0.16	0.16	0.12	0.10	0.10	0.12	0.10	0.10
Burundi	1.26	1.55	1.09	0.80	0.92	0.67	0.46	0.63	0.44	0.29
	0.11	0.12	0.11	0.11	0.06	0.05	0.04	0.06	0.05	0.04
Cameroon	6.22	6.38	5.24	5.03	2.66	2.61	2.23	1.98	2.07	1.72
	0.11	0.13	0.16	0.14	0.10	0.10	0.09	0.11	0.10	0.10
Chile	-0.44	-0.54	-0.54	-1.28	-2.65	-1.28	-1.70	-3.14	-1.36	-1.63
	0.11	0.11	0.11	0.13	0.10	0.09	0.08	0.11	0.10	0.08
Colombia	5.79	5.35	4.88	4.64	5.95	5.16	4.59	5.83	5.07	4.48
	0.09	0.10	0.10	0.10	0.08	0.07	0.07	0.08	0.08	0.07
Costa Rica	1.43	1.83	1.20	-0.08	4.39	6.06	6.10	2.96	5.11	5.09
	0.24	0.25	0.25	0.29	0.22	0.18	0.14	0.26	0.21	0.17
Dem Rep Congo	0.65	1.03	0.71	0.88	0.78	0.71	0.61	0.64	0.57	0.46
	0.07	0.08	0.07	0.09	0.06	0.06	0.05	0.07	0.07	0.06
Ecuador	6.22	6.09	5.93	5.90	2.92	3.60	3.47	2.70	3.49	3.34
	0.08	0.09	0.09	0.10	0.08	0.07	0.06	0.09	0.08	0.07
Mexico	8.81	9.50	7.93	7.09	4.35	4.85	5.80	2.46	3.22	4.18
	0.13	0.15	0.15	0.16	0.12	0.10	0.08	0.13	0.11	0.09
Morocco	12.80	13.49	9.22	12.55	4.62	4.36	0.36	4.05	3.87	0.10
	0.11	0.12	0.15	0.13	0.10	0.08	0.06	0.11	0.09	0.07
Mozambique	3.85	4.36	3.05	2.48	1.01	0.11	-0.60	0.21	-0.43	-0.96
	0.22	0.24	0.24	0.25	0.16	0.15	0.13	0.18	0.17	0.14
Niger	0.04	-0.17	0.20	-0.37	-0.77	-0.70	-0.85	-0.94	-0.81	-0.94
	0.20	0.21	0.24	0.24	0.19	0.18	0.16	0.22	0.21	0.18
Papua New Guinea	8.18	8.77	5.40	5.31	7.12	6.49	4.05	4.05	3.63	2.61
	0.33	0.33	0.32	0.33	0.27	0.26	0.20	0.27	0.26	0.21
Peru	5.49	5.63	5.21	4.95	3.96	3.96	2.26	3.73	3.73	2.23
	0.09	0.10	0.11	0.12	0.06	0.06	0.04	0.07	0.07	0.05
Rep of Congo	2.14	2.45	2.00	1.32	0.71	0.95	0.35	0.50	0.78	0.32
	0.17	0.18	0.18	0.19	0.12	0.09	0.09	0.13	0.10	0.09
Rwanda	2.50	2.49	1.98	2.33	0.04	0.02	0.00	0.03	0.00	-0.02
	0.09	0.10	0.10	0.11	0.02	0.02	0.02	0.02	0.02	0.02
South Africa	6.24	7.61	6.78	6.50	5.23	6.35	6.20	4.30	5.51	5.47
	0.09	0.11	0.11	0.12	0.09	0.08	0.07	0.09	0.09	0.07
Tanzania	1.98	1.80	1.76	1.26	1.35	1.15	0.58	1.17	0.78	0.22
	0.18	0.19	0.19	0.20	0.16	0.16	0.14	0.18	0.18	0.16
Uruguay	9.94	10.17	9.88	8.58	5.33	4.73	4.74	4.28	3.72	3.72
	0.47	0.51	0.46	0.57	0.23	0.21	0.19	0.22	0.19	0.17
All Countries (Mean)	4.47	4.71	3.82	3.65	2.57	2.78	2.16	1.90	2.23	1.73
Household Characteristics		X	X	X	X	X	X	X	X	X
Urban/Rural			X							
Minimum Geoloc				X				X	X	X
COICOP 2-dig					X			X		
COICOP 3-dig						X			X	
COICOP 4-dig							X			X

This table shows the average slope of the Informal Engel curve across countries for different specifications. In this robustness scenario, all stores that have been classified in category 3 (3: market consumption, small store-front) are assigned to the informal sector. The slopes are estimated from: $Share\,Informal_i = \beta.ln(expenditure\,pc)_i + \Gamma X_i + \varepsilon_i$, where the dependent variable is the informal expenditure share and the explanatory variable is the log expenditure pp. Controls include household characteristics (household size, age, gender, and education of head), geographic indicators (urban/rural and districts), and product codes at the 2nd, 3rd and 4th level of the United Nation's COICOP classification. While all countries follow the COICOP at the 2nd level (12 categories: food, clothing etc.), Brazil, Morocco and Peru have specific product classifications at lower levels. The geographic control "district" refers to the lowest geographical level available in each survey.

Table C3: Regression Results - Robustness Scenario C

	Main		Geography		Product Codes			District + Products		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	No control	Hhld Char.	Rur/Urb.	District	Level2	Level3	Level4	Level2	Level3	Level4
Brazil	1.92	2.02	1.73	1.38	1.38	3.05	3.05	0.80	2.59	2.64
	0.06	0.06	0.06	0.06	0.06	0.05	0.05	0.06	0.06	0.05
Burkina Faso	7.31	7.66	4.96	6.78	5.76	5.58	4.78	5.20	5.09	4.42
	0.22	0.24	0.23	0.23	0.17	0.17	0.15	0.17	0.16	0.15
Burundi	1.69	2.05	1.48	1.19	1.32	0.99	0.72	0.97	0.68	0.49
	0.12	0.13	0.12	0.12	0.07	0.05	0.04	0.07	0.06	0.05
Cameroon	6.65	6.86	4.98	5.43	4.33	4.21	3.46	3.40	3.42	2.60
	0.13	0.15	0.18	0.16	0.13	0.12	0.11	0.14	0.13	0.12
Chile	1.58	1.66	1.63	0.90	1.23	1.26	1.23	0.55	0.68	0.67
	0.06	0.07	0.06	0.07	0.06	0.05	0.05	0.06	0.06	0.05
Colombia	-0.39	-0.79	-0.94	-0.78	1.22	1.19	1.16	0.88	0.84	0.80
	0.05	0.06	0.06	0.06	0.05	0.05	0.04	0.05	0.05	0.04
Costa Rica	0.44	1.13	0.86	0.31	2.18	3.55	3.77	1.51	3.09	3.25
	0.18	0.20	0.20	0.24	0.17	0.16	0.14	0.21	0.19	0.17
Dem Rep Congo	1.64	2.85	2.07	2.82	2.09	1.76	1.49	2.10	1.76	1.52
	0.11	0.12	0.12	0.13	0.09	0.08	0.06	0.10	0.09	0.07
Ecuador	3.73	4.12	2.41	2.08	3.99	3.90	4.04	1.86	1.77	2.23
	0.12	0.13	0.13	0.13	0.12	0.12	0.09	0.12	0.12	0.10
Mexico	1.83	1.83	1.01	-0.34	0.68	1.36	2.55	-1.30	-0.47	0.82
	0.10	0.11	0.11	0.11	0.11	0.10	0.08	0.11	0.10	0.09
Morocco	12.11	12.39	5.25	11.09	7.18	7.52	5.25	6.22	6.59	4.56
	0.14	0.16	0.15	0.17	0.12	0.09	0.08	0.12	0.10	0.08
Mozambique	6.25	7.26	5.52	4.98	3.99	3.02	2.09	2.93	2.22	1.56
	0.25	0.27	0.27	0.27	0.20	0.19	0.16	0.20	0.19	0.16
Niger	1.34	1.27	0.07	0.88	1.50	1.48	1.75	1.13	1.10	1.40
	0.27	0.29	0.34	0.32	0.28	0.27	0.23	0.31	0.31	0.26
Papua New Guinea	7.73	8.31	4.62	4.36	6.84	6.02	2.78	3.31	2.80	1.48
	0.34	0.34	0.32	0.33	0.29	0.28	0.20	0.28	0.27	0.21
Peru	-0.17	-0.50	0.09	1.84	-1.67	-1.67	-1.71	0.77	0.77	0.86
	0.11	0.11	0.13	0.13	0.09	0.09	0.06	0.11	0.11	0.07
Rep of Congo	4.42	5.25	4.65	3.65	3.42	3.31	2.51	2.75	2.65	2.07
	0.21	0.22	0.22	0.23	0.15	0.13	0.11	0.16	0.14	0.12
Rwanda	2.93	2.82	0.98	1.10	0.87	0.49	1.73	-0.73	-1.16	0.51
	0.13	0.13	0.14	0.15	0.09	0.08	0.06	0.10	0.09	0.07
South Africa	1.39	1.84	1.60	1.69	1.55	1.83	1.84	1.44	1.79	1.81
	0.05	0.06	0.06	0.07	0.05	0.05	0.05	0.05	0.05	0.05
Tanzania	4.80	3.94	3.71	0.23	3.55	3.29	2.86	0.72	0.33	0.31
	0.28	0.29	0.29	0.27	0.24	0.23	0.19	0.24	0.23	0.20
Uruguay	0.82	0.81	0.78	0.88	0.47	0.39	0.44	0.58	0.50	0.54
	0.06	0.07	0.08	0.09	0.09	0.08	0.08	0.10	0.09	0.09
All Countries (Mean)	3.40	3.64	2.37	2.52	2.59	2.63	2.29	1.75	1.85	1.73
Household Characteristics		X	X	X	X	X	X	X	X	X
Urban/Rural			X							
Minimum Geoloc				X				X	X	X
COICOP 2-dig					X			X		
COICOP 3-dig						X			X	
COICOP 4-dig							X			X

This table shows the average slope of the Informal Engel curve across countries for different specifications. In this robustness scenario, all stores that have been classified in category 3 (3: market consumption, small store-front) are assigned to the formal sector. The slopes are estimated from: $Share\,Informal_i = \beta.ln(expenditure\,pc)_i + \Gamma X_i + \varepsilon_i$, where the dependent variable is the informal expenditure share and the explanatory variable is the log expenditure pp. Controls include household characteristics (household size, age, gender, and education of head), geographic indicators (urban/rural and districts), and product codes at the 2nd, 3rd and 4th level of the United Nation's COICOP classification. While all countries follow the COICOP at the 2nd level (12 categories: food, clothing etc.), Brazil, Morocco and Peru have specific product classifications at lower levels. The geographic control "district" refers to the lowest geographical level available in each survey.

Table C4: Regression Results

	Main		Geography			Product Codes			District + Products		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	No control	Hhld Char.	Rur/Urb.	District	CensusBlock	Level2	Level3	Level4	Level2	Level3	Level4
Brazil	1.92	2.02	1.73	1.38	0.90	1.38	3.05	3.05	0.80	2.59	2.64
	0.06	0.06	0.06	0.06	0.07	0.06	0.05	0.05	0.06	0.06	0.05
Burkina Faso	5.70	6.01	4.03	5.47	3.91	3.68	3.47	2.41	3.52	3.30	2.34
	0.18	0.19	0.18	0.19	0.20	0.12	0.11	0.11	0.12	0.11	0.11
Burundi	1.69	2.05	1.48	1.19	0.63	1.32	0.99	0.72	0.97	0.68	0.49
	0.12	0.13	0.12	0.12	0.09	0.07	0.05	0.04	0.07	0.06	0.05
Cameroon	6.85	7.13	5.99	5.81	4.88	3.17	2.98	2.52	2.51	2.44	2.01
	0.12	0.13	0.17	0.14	0.21	0.11	0.10	0.10	0.12	0.11	0.10
Chile	0.61	0.55	0.55	0.20	0.05	-1.86	-1.98	-2.54	-1.92	-1.95	-2.37
	0.09	0.10	0.10	0.11	0.12	0.08	0.07	0.06	0.09	0.08	0.06
Colombia	7.11	6.75	6.36	5.94	5.08	6.60	6.21	5.44	6.26	5.89	5.15
	0.09	0.10	0.10	0.10	0.11	0.07	0.07	0.06	0.07	0.07	0.06
Costa Rica	5.04	5.99	5.33	3.93	3.93	6.20	7.79	7.88	4.57	6.38	6.46
	0.22	0.24	0.24	0.28	0.28	0.20	0.19	0.15	0.24	0.22	0.18
Dem Rep Congo	1.64	2.85	2.07	2.82	2.95	2.09	1.76	1.49	2.10	1.76	1.52
	0.11	0.12	0.12	0.13	0.17	0.09	0.08	0.06	0.10	0.09	0.07
Ecuador	10.82	10.80	10.44	10.08	9.63	6.85	6.84	6.77	5.98	5.99	6.13
	0.10	0.11	0.11	0.12	0.12	0.09	0.09	0.08	0.10	0.10	0.09
Mexico	9.01	9.90	8.29	6.87	6.87	5.83	6.34	6.52	3.14	3.69	4.11
	0.12	0.14	0.14	0.15	0.15	0.12	0.11	0.09	0.13	0.12	0.10
Morocco	14.34	15.27	10.40	14.17	8.21	6.99	6.19	2.47	6.31	5.53	2.03
	0.11	0.12	0.14	0.13	0.17	0.10	0.08	0.06	0.10	0.09	0.07
Mozambique	6.25	7.26	5.52	4.98	5.42	3.99	3.02	2.09	2.93	2.22	1.56
	0.25	0.27	0.27	0.27	0.28	0.20	0.19	0.16	0.20	0.19	0.16
Niger	0.07	-0.13	0.25	-0.32	-0.29	-0.74	-0.67	-0.83	-0.89	-0.77	-0.91
	0.20	0.21	0.25	0.24	0.26	0.19	0.18	0.16	0.22	0.21	0.18
Papua New Guinea	8.18	8.77	5.40	5.31	7.19	7.12	6.49	4.05	4.05	3.63	2.61
	0.33	0.33	0.32	0.33	0.35	0.27	0.26	0.20	0.27	0.26	0.21
Peru	7.22	7.59	7.02	7.16	6.12	5.40	5.40	4.25	4.62	4.62	3.72
	0.10	0.10	0.12	0.13	0.15	0.06	0.06	0.05	0.08	0.08	0.06
Rep of Congo	4.42	5.25	4.65	3.65	6.33	3.42	3.31	2.51	2.75	2.65	2.07
	0.21	0.22	0.22	0.23	0.29	0.15	0.13	0.11	0.16	0.14	0.12
Rwanda	3.55	3.53	2.78	3.11	2.72	1.20	0.98	0.93	0.94	0.66	0.63
	0.11	0.11	0.11	0.12	0.13	0.06	0.05	0.05	0.06	0.05	0.05
South Africa	1.39	1.84	1.60	1.69		1.55	1.83	1.84	1.44	1.79	1.81
	0.05	0.06	0.06	0.07		0.05	0.05	0.05	0.05	0.05	0.05
Tanzania	4.80	3.94	3.71	0.23	2.14	3.55	3.29	2.86	0.72	0.33	0.31
	0.28	0.29	0.29	0.27	0.29	0.24	0.23	0.19	0.24	0.23	0.20
Uruguay	8.18	8.47	8.22	7.44	7.31	5.12	4.91	4.89	4.34	4.13	4.04
	0.17	0.19	0.19	0.22	0.22	0.23	0.21	0.16	0.22	0.19	0.14
All Countries (Mean)	5.44	5.79	4.79	4.56	4.29	3.64	3.61	2.97	2.76	2.78	2.32

Table C5: Regression Results: Robustness Scenario B

	Main		Geography			Product Codes			District + Products		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	No control	Hhld Char.	Rur/Urb.	District	CensusBlock	Level2	Level3	Level4	Level2	Level3	Level4
Brazil	2.13	2.03	1.62	1.23	0.22	0.78	3.46	2.79	0.00	2.92	2.33
	0.07	0.07	0.07	0.07	0.08	0.07	0.05	0.04	0.07	0.05	0.04
Burkina Faso	4.08	4.29	2.89	3.92	2.73	2.60	2.38	1.67	2.55	2.31	1.68
	0.15	0.16	0.16	0.16	0.17	0.12	0.10	0.10	0.12	0.10	0.10
Burundi	1.26	1.55	1.09	0.80	0.31	0.92	0.67	0.46	0.63	0.44	0.29
	0.11	0.12	0.11	0.11	0.08	0.06	0.05	0.04	0.06	0.05	0.04
Cameroon	6.22	6.38	5.24	5.03	4.14	2.66	2.61	2.23	1.98	2.07	1.72
	0.11	0.13	0.16	0.14	0.21	0.10	0.10	0.09	0.11	0.10	0.10
Chile	-0.44	-0.54	-0.54	-1.28	-1.68	-2.65	-1.28	-1.70	-3.14	-1.36	-1.63
	0.11	0.11	0.11	0.13	0.14	0.10	0.09	0.08	0.11	0.10	0.08
Colombia	5.79	5.35	4.88	4.64	3.71	5.95	5.16	4.59	5.83	5.07	4.48
	0.09	0.10	0.10	0.10	0.11	0.08	0.07	0.07	0.08	0.08	0.07
Costa Rica	1.43	1.83	1.20	-0.08	-0.08	4.39	6.06	6.10	2.96	5.11	5.09
	0.24	0.25	0.25	0.29	0.29	0.22	0.18	0.14	0.26	0.21	0.17
Dem Rep Congo	0.65	1.03	0.71	0.88	0.98	0.78	0.71	0.61	0.64	0.57	0.46
	0.07	0.08	0.07	0.09	0.12	0.06	0.06	0.05	0.07	0.07	0.06
Ecuador	6.22	6.09	5.93	5.90	5.45	2.92	3.60	3.47	2.70	3.49	3.34
	0.08	0.09	0.09	0.10	0.10	0.08	0.07	0.06	0.09	0.08	0.07
Mexico	8.81	9.50	7.93	7.09	7.09	4.35	4.85	5.80	2.46	3.22	4.18
	0.13	0.15	0.15	0.16	0.16	0.12	0.10	0.08	0.13	0.11	0.09
Morocco	12.80	13.49	9.22	12.55	6.99	4.62	4.36	0.36	4.05	3.87	0.10
	0.11	0.12	0.15	0.13	0.19	0.10	0.08	0.06	0.11	0.09	0.07
Mozambique	3.85	4.36	3.05	2.48	2.72	1.01	0.11	-0.60	0.21	-0.43	-0.96
	0.22	0.24	0.24	0.25	0.27	0.16	0.15	0.13	0.18	0.17	0.14
Niger	0.04	-0.17	0.20	-0.37	-0.37	-0.77	-0.70	-0.85	-0.94	-0.81	-0.94
	0.20	0.21	0.24	0.24	0.26	0.19	0.18	0.16	0.22	0.21	0.18
Papua New Guinea	8.18	8.77	5.40	5.31	7.19	7.12	6.49	4.05	4.05	3.63	2.61
	0.33	0.33	0.32	0.33	0.35	0.27	0.26	0.20	0.27	0.26	0.21
Peru	5.49	5.63	5.21	4.95	3.87	3.96	3.96	2.26	3.73	3.73	2.23
	0.09	0.10	0.11	0.12	0.14	0.06	0.06	0.04	0.07	0.07	0.05
Rep of Congo	2.14	2.45	2.00	1.32	3.48	0.71	0.95	0.35	0.50	0.78	0.32
	0.17	0.18	0.18	0.19	0.25	0.12	0.09	0.09	0.13	0.10	0.09
Rwanda	2.50	2.49	1.98	2.33	2.00	0.04	0.02	0.00	0.03	0.00	-0.02
	0.09	0.10	0.10	0.11	0.12	0.02	0.02	0.02	0.02	0.02	0.02
South Africa	6.24	7.61	6.78	6.50		5.23	6.35	6.20	4.30	5.51	5.47
	0.09	0.11	0.11	0.12		0.09	0.08	0.07	0.09	0.09	0.07
Tanzania	1.98	1.80	1.76	1.26	1.41	1.35	1.15	0.58	1.17	0.78	0.22
	0.18	0.19	0.19	0.20	0.19	0.16	0.16	0.14	0.18	0.18	0.16
Uruguay	9.94	10.17	9.88	8.58	8.51	5.33	4.73	4.74	4.28	3.72	3.72
	0.47	0.51	0.46	0.57	0.55	0.23	0.21	0.19	0.22	0.19	0.17
All Countries (Mean)	4.47	4.71	3.82	3.65	3.31	2.57	2.78	2.16	1.90	2.23	1.73

Table C6: Regression Results: Robustness Scenario C

	Main		Geography			Product Codes			District + Products		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	No control	Hhld Char.	Rur/Urb.	District	CensusBlock	Level2	Level3	Level4	Level2	Level3	Level4
Brazil	1.92	2.02	1.74	1.39	0.91	1.39	3.07	3.07	0.81	2.62	2.66
	0.06	0.06	0.06	0.06	0.07	0.06	0.05	0.05	0.06	0.06	0.05
Burkina Faso	7.31	7.61	4.90	6.84	4.72	5.68	5.55	4.77	5.22	5.15	4.48
	0.22	0.25	0.25	0.24	0.25	0.19	0.18	0.15	0.18	0.17	0.15
Burundi	1.69	2.11	1.54	1.26	0.68	1.36	1.02	0.75	1.01	0.71	0.52
	0.12	0.13	0.12	0.12	0.10	0.07	0.06	0.05	0.07	0.06	0.05
Cameroon	6.65	7.03	5.05	5.51	2.97	4.55	4.44	3.70	3.58	3.64	2.80
	0.13	0.15	0.18	0.16	0.23	0.13	0.13	0.11	0.14	0.14	0.12
Chile	1.58	1.67	1.64	0.90	0.73	1.23	1.27	1.23	0.55	0.68	0.68
	0.06	0.07	0.06	0.07	0.07	0.06	0.05	0.05	0.06	0.06	0.05
Colombia	-0.39	-0.79	-0.94	-0.77	-1.10	1.23	1.20	1.18	0.90	0.86	0.82
	0.05	0.06	0.06	0.06	0.07	0.05	0.05	0.04	0.05	0.05	0.04
Costa Rica	0.44	1.19	0.93	0.42	0.42	2.25	3.63	3.86	1.61	3.22	3.39
	0.18	0.20	0.20	0.24	0.24	0.18	0.16	0.15	0.21	0.19	0.17
Dem Rep Congo	1.64	3.02	2.22	3.06	3.53	2.18	1.82	1.54	2.22	1.86	1.60
	0.11	0.13	0.12	0.14	0.19	0.09	0.08	0.07	0.10	0.09	0.08
Ecuador	3.73	4.07	2.32	1.97	1.83	4.02	3.94	4.07	1.89	1.80	2.28
	0.12	0.13	0.13	0.13	0.14	0.12	0.12	0.09	0.12	0.12	0.10
Mexico	1.83	1.85	1.03	-0.35	-0.35	0.69	1.38	2.58	-1.33	-0.48	0.83
	0.10	0.11	0.12	0.12	0.12	0.11	0.10	0.08	0.11	0.11	0.09
Morocco	12.11	12.52	5.29	11.25	3.07	7.26	7.61	5.32	6.29	6.69	4.63
	0.14	0.16	0.16	0.17	0.17	0.12	0.09	0.08	0.13	0.10	0.08
Mozambique	6.25	7.28	5.52	5.09	5.56	3.98	3.02	2.05	3.00	2.31	1.62
	0.25	0.28	0.28	0.28	0.29	0.20	0.19	0.16	0.20	0.20	0.17
Niger	1.34	1.33	0.11	0.98	-0.33	1.55	1.53	1.80	1.20	1.19	1.47
	0.27	0.29	0.35	0.33	0.39	0.28	0.28	0.23	0.32	0.31	0.26
Papua New Guinea	7.73	8.43	4.72	4.45	6.90	6.94	6.12	2.84	3.37	2.87	1.52
	0.34	0.34	0.32	0.33	0.34	0.29	0.28	0.20	0.28	0.27	0.21
Peru	-0.17	-0.42	0.11	1.85	1.14	-1.61	-1.61	-1.67	0.78	0.78	0.87
	0.11	0.11	0.13	0.13	0.14	0.09	0.09	0.06	0.11	0.11	0.07
Rep of Congo	4.42	5.19	4.61	3.58	6.48	3.53	3.42	2.60	2.86	2.77	2.16
	0.21	0.22	0.22	0.23	0.30	0.15	0.13	0.11	0.16	0.14	0.12
Rwanda	2.93	2.82	0.98	1.10	0.27	0.87	0.49	1.73	-0.73	-1.16	0.51
	0.13	0.13	0.14	0.15	0.15	0.09	0.08	0.06	0.10	0.09	0.07
South Africa	1.39	1.84	1.60	1.69		1.55	1.83	1.84	1.44	1.79	1.81
	0.05	0.06	0.06	0.07		0.05	0.05	0.05	0.05	0.05	0.05
Tanzania	4.80	3.94	3.71	0.23	2.14	3.55	3.29	2.86	0.72	0.33	0.31
	0.28	0.29	0.29	0.27	0.29	0.24	0.23	0.19	0.24	0.23	0.20
Uruguay	0.82	0.81	0.78	0.88	0.84	0.47	0.39	0.44	0.58	0.50	0.54
	0.06	0.07	0.08	0.09	0.09	0.09	0.08	0.08	0.10	0.09	0.09
All Countries (Mean)	3.40	3.68	2.93	2.57	2.11	2.63	2.67	2.33	1.80	1.91	1.78