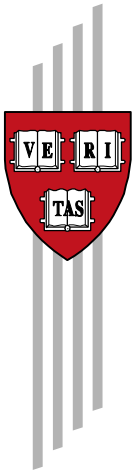
Informality, Consumption Taxes and Redistribution

Pierre Bachas, Lucie Gadenne & Anders Jensen

CID Faculty Working Paper No. 407
September 2021



Working Papers
Center for International Development
at Harvard University

Informality, Consumption Taxes and Redistribution

Pierre Bachas, Lucie Gadenne & Anders Jensen*

September 2021

Abstract

Can taxes on consumption redistribute in developing countries? Contrary to consensus, we show that taxing consumption is progressive once we account for informal consumption. Using household expenditure surveys in 32 countries we proxy for informal consumption using the type of store where purchases occur. We find that the budget share spent in informal stores steeply declines with income, so that the effective tax rate of a broad consumption tax rises with income. Our findings imply that the widespread policy of exempting food from taxation cannot be justified on equity grounds in low-income-countries.

JEL: E26, H21, H23, 023

Keywords: Budget Surveys, Inequality, Informality, Redistribution, Taxes.

*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. We would like to thank Michael Best, Anne Brockmeyer, Roberto Fatal, Rema Hanna, Gordon Hanson, Xavier Jaravel, Michael Keen, Joana Naritomi, Henrik Kleven, Wojciech Kopczuk, Joel Slemrod, Johannes Spinnewijn, Mazhar Waseem and numerous seminar participants for helpful comments. We thank Eva Davoine, Elie Gerschel, Mariana Racimo, Roxanne Rahnama and Alvaro Zuniga for excellent research assistance. We gratefully acknowledge financial support from the Weatherhead Center for International Affairs at Harvard University, the World Bank Development Economics' Research Support Budget, and UKAID through the IFS's TaxDev Centre. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. Replication codes for the paper are available here.

1 Introduction

Income inequality in developing countries is high and has persisted over the past 30 years (Alvaredo and Gasparini, 2015). Direct income taxes are constrained in these countries (Jensen, 2019), making indirect taxes the main source of government revenue. What is the redistributive impact of taxes on consumption? Since uniform consumption taxes are viewed as inequitable, 90% of developing countries apply reduced rates on necessity goods, in particular food items. Do rate-differentiation policies reduce inequality? In this paper, we answer these questions by systematically investigating consumption taxes' redistributive capacity in developing countries. Our focus on the equity characteristics of taxes builds upon and departs from the literature in public finance and development which focuses on their revenue and efficiency properties (Besley and Persson, 2013).

Using a large micro database across countries and a new method to proxy for household consumption from the informal sector, our analysis yields two main findings. First, due to differences in informal consumption along the income distribution, uniform consumption taxes are progressive and reduce inequality. This result runs counter to the consensus view that taxes on consumption have negative or neutral distributional impacts. Second, the redistributive impact of rate differentiation is severely weakened when accounting for informal consumption; in particular, our results imply that the widespread policy of exempting food from taxation cannot be justified on equity grounds in low-income countries.

Our starting point is the construction of a micro database of expenditure surveys from 32 low and middle income countries. We innovate by using the store type reported for each purchase in those surveys to proxy for household consumption from the informal sector. This approach is motivated by the vast disparities in consumption by place of purchase across countries: Figure 1a-1b shows that economic development is associated with a rise in consumption in modern stores (supermarkets, specialized stores), which gradually replaces consumption in traditional 'stores' (home production, street stalls, corner stores). Modern and traditional stores differ in structural characteristics which determine tax enforceability, including size, organizational structure and interaction with third-parties. In our main formality assignment, we assume that taxes are remitted on purchases from modern stores and not from traditional stores. We provide both descriptive

evidence across countries and quasi-experimental evidence within country suggesting that this stylized formality assignment is reasonable to a first order.

We use this database to establish new stylized facts on consumption patterns across the household income distribution and between countries. We document the existence of a downward sloping Informality Engel Curve (IEC): the informal budget share steeply declines with household income in every country. Our data also allows us to study patterns of informal consumption within goods. We focus on food versus non-food since most countries tax food at a reduced rate. Accounting for informal consumption fundamentally alters the goods level patterns: while the overall food Engel curve is steep and negative in all countries, the formal food Engel curve has a small but *positive* slope in low-income countries, and only becomes negative in upper-middle income countries.

These patterns determine the progressivity of consumption taxes. A tax is progressive if the effective tax rate (ratio of taxes paid to household income) increases with income. In our average country, a uniform tax rate levied on all formal consumption is strongly progressive due to the downward-sloping IECs: the effective tax rate of the richest quintile of households is twice that of the poorest. Moreover, the progressivity gain from exempting food while taxing non-food is limited since poor households' food consumption mostly occurs in informal stores.

The progressivity of consumption tax policies varies across countries. The progressivity achieved with a uniform rate decreases with development. This is because the aggregate informal budget share is large in low-income countries, which, combined with a negative IEC slope, implies that a formal purchase is a strong tag for high income households (Akerlof, 1978). As the informal budget share shrinks with development, formal purchases become a weaker income tag. On the contrary, exempting food from taxation produces no progressivity gains in the poorest countries, but leads to moderate gains in upper-middle income countries. Failing to account for informality leads to an overestimation of progressivity gains from food exemption in all countries, but particularly in the poorest ones.

What are the implications of these consumption patterns for tax design? To study this we extend Diamond (1975)'s multi-person model of optimal commodity taxes to allow for formal and informal (untaxable) varieties of each good. The model enables us to take into account both the equity and efficiency implications of informal consumption: introducing informal varieties increases the efficiency

cost of taxes since households substitute consumption towards them when taxes increases. This cost decreases over development as the informal sector shrinks (Figure 1a). Calibrating the model to our data, we find that the optimal level of rate differentiation between food and non-food increases with development. In low-income countries, reduced rates often cannot be justified on equity grounds.

To measure the impact of consumption taxes on inequality we combine the calibrated rates with our microdata. We find that setting optimal uniform rates reduces the Gini coefficient by 1.9% on average: 1% in low-income countries and up to 3% in upper middle-income countries; with rate differentiation, the inequality reduction ranges from 1.1% to 3.9%.[1] We compare our results to the findings from Commitment to Equity (CEQ), which evaluates actual policies and whose findings reflect the consensus view on the redistributive role of consumption taxes (Lustig, 2018). Indeed, the average Gini reduction from consumption taxes in CEQ across 25 developing countries is 0.6% – three times smaller than our average estimate. Our results are more comparable in magnitude to the average inequality achieved by direct income taxes in CEQ (2.6%). In an extension, we incorporate an income tax into our model. While the presence of a perfectly enforceable income tax renders redistribution through consumption taxes suboptimal (Atkinson and Stiglitz, 1976), we find that taking into account the imperfectly enforced income taxes in place in developing countries lowers the inequality reduction achievable via consumption taxes by 10%.

Our main assignment is based on a stylized incidence assumption where all modern (traditional) stores are formal (informal) with 100% (0%) pass-through of taxes to consumer prices. We combine micro-data on firms by store type and formality status in our sample countries and estimate that 85% of modern stores are formal, compared to 10% of traditional stores. Moreover, formality shares by store type are constant across low and middle income countries. To gauge our assumption of differential pass-through by store type, we directly estimate pass-through in modern and traditional stores in Mexico, one of our sample countries, where a reform increased the consumption tax rate only in some locations. We find a 14% pass-through in traditional stores (not statistically significant) and a 77% pass-through in modern stores. Applying these estimates to every country

[1] Our main results focus on food versus non-food rate differentiation, but we also show results for optimal differentiation between 12 large goods categories.

reduces the inequality gains of consumption taxes but leaves our results qualitatively unchanged. We also discuss how pass-through may differ in other countries and in more complex incidence settings (Benzarti and Carloni, 2019). More generally, our results hold qualitatively for *any* positive difference in pass-through between modern and traditional stores. As more pass-through estimates by store type become available, they can be combined with our model, method, and publicly available data to assess inequality effects.[2]

Our paper provides two main contributions. First, using a novel methodology and dataset, we show that consumption taxes are progressive and reduce inequality in developing countries. This finding runs counter to the consensus view which argues that indirect taxes have no redistributive potential. This consensus is based on limited empirical evidence which exists on an ad-hoc country-by-country basis and typically ignores the role played by informal consumption (Sah, 1983; Gemmell and Morrissey, 2005; Harris et al., 2018). Our paper presents a new distributional method to study the equity of consumption taxes, the main tax base in developing countries.[3] Except for recent work on wealth taxes (Londono-Velez and Avila-Mahecha, 2021), studies on tax and development rarely analyze equity consequences of imperfect enforcement, but focus on revenue and efficiency implications (Bergeron et al., 2021; Best et al., 2015; Kleven and Waseem, 2013).

Second, we conduct the first systematic analysis of the redistributive potential of optimal consumption tax rate differentiation.[4] Our results show that it depends on how informal consumption along the income distribution alters the shape of goods' Engel curves. This relates to recent papers on the equity implications of distributional consumption patterns in rich countries (Faber and Fally, 2017; Jaravel, 2019; Allcott et al., 2019). We find that exempting necessities only modestly redistributes once we account for informal consumption, especially in low-income countries. This result contributes to the literature on the design of indirect taxes in developing countries (Pomeranz, 2015; Naritomi, 2019; Waseem, 2020).

[2]Estimating pass-through by store type in every country would require micro-level price quotes separately for modern vs traditional stores, and VAT reforms providing credible research designs in 32 countries. Instead, we provide a model-based discussion, combined with descriptive evidence, to gauge how the pass-through may vary across countries.

[3]Recent studies also find societal benefits from increased revenue and improved tax capacity (Casaburi and Troiano, 2015; Gadenne, 2017; Weigel, 2019; De Simone, 2020).

[4]Country reports from the International Bureau of Fiscal Documentation show that 90% of developing countries differentiate tax rates, in particular between food and non-food goods.

Finally, our new micro database allows us to document how consumption by place of purchase varies with household income both within countries and across development, thus relating to macro-studies which focus on aggregate changes in consumption by store type across countries (Bronnenberg and Ellickson, 2015; Lagakos, 2016; Atkin et al., 2018b). Our methodology allows us to construct a new consumption based measure of informality, which complements pre-existing work that focuses on informality at the firm and worker level (La Porta and Shleifer, 2014; Gerard and Gonzaga, 2016; Ulyssea, 2018).

2 Data

In this section, we describe the data sources and selection criteria used to construct our micro dataset. Next, we outline how we measure consumption by store type. Finally, we describe secondary datasets.

2.1 Data Sources and Selection Criteria

We assemble our main dataset by combining household expenditure surveys from developing countries that satisfy three selection criteria. First, the survey must be nationally representative. Second, the survey must record consumption from open diaries rather than pre-filled diaries, which only contain information on selected goods. This helps to ensure that the survey covers all expenditure types. Third, the diary must ask households to report the store type where each item is purchased - the *place of purchase* - and this information must be systematically reported in the diaries. This last criterion ensures that we can apply our method to infer consumption from informal sources, described below.

Using these criteria, we includes surveys from 32 countries covering approximately 400,000 households. Table 1 lists alphabetically the countries in the data, with their survey name and year, the number of households, and the average number of purchases reported per household. Countries in the sample are principally located in Latin America and Sub-Saharan Africa. Unfortunately, most household expenditure surveys in Asia do not contain information on the place of purchase. Nonetheless, our dataset covers a wide range of income levels, from Burundi to Chile. Appendix B provides further details on the data sources used.

2.2 Measuring Consumption by Store Type

Our objective is to measure consumption by place of purchase so that it can be compared between households within and across countries. A challenge is that the 32 surveys do not share the same design. We create a taxonomy of place of purchases which aims to achieve international comparability, drawing on the framework established by the International Price Comparison Program. Our taxonomy contains seven categories for place of purchase. The first five pertain to purchases of goods: (1) non-market consumption (e.g. home production); (2) non brick and mortar stores (e.g. street stalls, public markets); (3) corner and convenience stores; (4) specialized stores (e.g. clothing stores); and, (5) large stores (e.g. supermarkets, department stores). Purchases of services are allocated to two main categories: (6) services provided by an institution (e.g. banks, hospitals); and, (7) services provided by an individual (e.g. domestic services).[5] These categories account for 86% of total household expenditure. The remaining 14% are items for which no place of purchase is specified, primarily utilities, fuel and telecommunication (see Figures A1 and E2).

We use an aggregated store classification for our main analysis, assigning categories (1) through (3) to the *traditional store type*, and categories (4) and (5) to the *modern store type*.[6] We do this for two reasons. First, the modern-traditional classification is commonly used in cross-country academic studies (Reardon et al., 2003; Humphrey, 2007; Lagakos, 2016) and market research on global retail patterns. It is based on the logic that differences in retailing across space and time are captured meaningfully by focusing on these two retail groups, since store types within each group share similar characteristics in most settings but are systematically different across groups in terms of sales, market orientation, and organizational structure. Second, as discussed in Section 3, these store types differ significantly in tax enforceability characteristics and compliance status.

Finally, we classify goods according to the UN's COICOP methodology. This allows us to observe how purchases in modern and traditional stores differ within increasingly narrow product categories; we focus on food vs non-food, as well as the 12/47/117 goods categories of the COICOP 2-digit/3-digit/4-digit level.

[5] All original store types in all surveys, and our classification, is shown in the online appendix. We exclude housing expenditure due to limited data on imputed rents.

[6] We explain how we assign services when we discuss the tax status by store type in Section 3.

2.3 Additional Data

Euromonitor market research We use data from the country specific retail reports produced by the private market research firm Euromonitor International. These reports contain information on modern and traditional retail food stores for each country (N=189), including number of outlets and total sales. The modern and traditional categories are consistent with our classification, with the exception that Euromonitor does not measure home-based consumption. The data reported is based on direct collection from retailers, surveys of retail trade, desk research and public data sources (see Appendix C.1 and Bronnenberg and Ellickson, 2015).

Mexican retail census and prices We use two datasets from Mexico. First, the 2013 Census collects information on the universe of retailers, including taxes levied on sales and paid on inputs. Second, we use the confidential monthly price quotes collected by the statistics office. This data samples prices for all items, stores, and locations representative of Mexican consumption. Importantly, both datasets contain details on store types that are consistent with our cross-country taxonomy (Section 2.2), with the caveat that home-based consumption is not included (details in Appendix D.4).

3 Measurement of Informal Consumption

In this section, we describe the characteristics of modern and traditional retailers which determine tax enforceability. Next, we introduce our assignment of tax formality status by store type and provide supporting evidence.

3.1 Characteristics of Modern & Traditional Retailers

Modern and traditional stores differ in characteristics which are key determinants of tax enforceability. Figures 1c-1d show that, in most countries, the average modern store is forty times larger in sales than the average traditional store. This difference translates into enforcement intensity since tax administrations devote more resources to monitor larger firms (Basri et al., 2019). In addition, modern stores occupy twenty times more floor space than traditional stores (Figures 1e-1f). The extra space allows modern stores to accommodate more customers,

employ more personnel, and hold inventory and accounting records. Studies on global retail identify the adoption of advanced accounting records as a key driver of modern stores' expansion (Evenson, 2007). The public finance literature convincingly shows that information trails, through accounting records and reports by third-parties including customers, employees, and suppliers, are key determinants of tax enforcement success (Kleven, 2014; Pomeranz, 2015; Naritomi, 2019). Modern stores' size and third-party information coverage implies that they are much more likely to be tax compliant than traditional stores.[7]

3.2 Assignment of Formality by Store Type

Baseline assignment Our definition of formality is based on the likelihood of consumption taxes being levied on consumer prices in a particular store type. Motivated by the previous subsection, our baseline assignment considers that all purchases made in traditional stores are informal (categories 1 to 3) and all purchases from modern stores are formal (categories 4 and 5). For services, we assume that institutions (category 6) are formal while individual providers are informal (category 7).[8]

Within country evidence The baseline assignment presents the advantage of relying on an observable characteristic which is comparable across countries. While the store type is an ex-ante characteristic that captures potential formality status, we show that it strongly correlates with ex-post actual formality. The Mexican retail census collects information on consumption tax (VAT) payments for all retailers: we find that only 9.5% of traditional stores report remitting these taxes, whereas most modern stores do (Figure C1).

To our knowledge, censuses in other countries do not contain information on both tax status and store types. To measure formality status by store type across countries, we instead rely on the World Bank Enterprise and Informal Surveys (WBEIS) and the Euromonitor retail reports. The WBEIS surveys measure tax

[7] This is consistent with macro studies of retailing that assume traditional (modern) stores are evading (compliant), including Lagakos (2016).

[8] With the same logic we unspecified expenditures are assigned to formal retail, since they mainly consist of utilities provided by large firms which cannot evade taxes (Figure A1). In the online appendix., we show for each country the original names of the places of purchase, their expenditure shares and our formality assignment.

registration and sales of retailers, but not store type, in 35 developing countries. The retail reports allow us to compute the average sales of modern and traditional retailers in the same countries. We measure the formality share in modern stores as the intersection between the WBEIS sales distribution of formality share and the Euromonitor average sales of modern stores. We repeat this exercise to estimate the formal share of traditional stores.[9] We find that the formal share in modern stores is on average 80-90%; the formality share in traditional stores is 10% (Figure C2). Moreover, the formality share within each store type is fairly constant across countries. This suggests that our country-invariant baseline assignment of formality status to store type may be reasonable to a first order.

The stability of the formality of traditional and modern stores across countries is consistent with the view that the increased consumption tax base over development is mainly driven by the growth of modern retailing rather than by changes in enforceability within store type.[10]

Tax exemption thresholds The tax status of a store is in part driven by the extent of enforcement on store-owners that are legally obligated to comply with consumption taxes but try to avoid it. In addition, stores can be legally exempt from taxes if their size falls below the exemption threshold. We code the value of the consumption tax threshold in all sample countries and find that the ratio of average sales to the threshold is 1.01 for traditional stores and 38.85 for modern stores (Appendix C.1).[11] This suggests that the large differences in formality share between store types may occur in part as a result of the tax code: given their size relative to the exemption threshold, a significant share of traditional stores are informal because they are not legally required to remit taxes, while the large size

[9] Details in Appendix C.1, which also discusses issues from combining datasets. The WBEIS surveys and formality variable are the same as in La Porta and Shleifer (2014).

[10] Figure A2 shows the budget shares by detailed store categories within modern and traditional. Home-production and street stalls (categories 1 and 2) account for over 75% of consumption in traditional stores. There is arguably less uncertainty about their formality status than in the remaining traditional stores (corner stores, category 3).

[11] We use the Value-Added Tax (VAT) threshold since all countries in our sample use a VAT. The level of the exemption threshold is itself endogenous to enforcement constraints: the tax administration knows that the tax revenue-yield from taxing small stores can be minor relative to administrative and compliance costs, and thus chooses to exempt them (See Ebrill and Keen, 2001; Keen and Mintz, 2004).

of modern stores compels most of them to remit taxes.[12]

Summary Our baseline assignment of modern stores as formal and traditional stores as informal appears reasonable given the descriptive evidence. This assignment is transparent and constitutes our starting point, but we show below that results are robust to using the country-specific formality shares by store types obtained from the surveys and reports. Further we assume in what follows that consumption taxes are fully passed through to prices in formal stores and not at all in informal stores. In Section 8, we discuss this assumption in detail, provide quasi-experimental evidence on it using a Value-Added-Tax (VAT) reform in Mexico, and present robustness results with different pass-through assumptions.

4 Engel Curves of Informality and Food Across Development

In this section, we show how informal consumption varies with household income within and between countries, and investigate the determinants of these variations. We then document how food and non food consumption differ across stores, and how these patterns vary with income.

4.1 Informality Engel Curves

To study how informal consumption varies with income, we measure the informality Engel curve (IEC). The IEC traces the relationship between the informal budget share and total household expenditure within a country. We proxy income with total expenditure due to known issues with measuring income in developing countries (Deaton and Paxson, 1998; Atkin et al., 2018a). We use the logarithm of total household expenditure per person, in line with the literature on Engel curves (Deaton, 1997). For illustrative purposes, Figure 2 plots the IEC for a low-income country (Rwanda) and a middle-income country (Mexico). To investigate the functional form flexibly, the non-parametric IEC is constructed from local polynomial regressions. In Rwanda, the informal budget share falls from 90% for the poorest decile of households to 70% for the richest decile. In

[12]Traditional retailers below the exemption threshold may still voluntarily register to pay VAT. However, Almunia et al. (2019) predict that optimal voluntary registration is more prevalent for manufacturing firms than retailers, due to their location at intermediate levels in supply chains.

Mexico, the IEC falls from 55% to 25%. We find two empirical regularities in the full sample.[13] First, IECs slope downward everywhere. Second, IECs are approximately linear in log expenditure. This suggests a stable functional form relation between informal budget shares and household expenditure.[14]

We summarize the information contained in the country-level IECs with two empirical moments: i) the aggregate informal budget share; ii) the slope of the IEC. In Section 5, we explain how these two moments are sufficient to characterize the tax progressivity impacts of consumption patterns. In Figure 3, we plot the aggregate informal budget share (Figure 3a) and the estimated IEC slope (Figure 3b) against countries' GDP per capita. Figure 3a reveals a large drop in the aggregate informal budget share, from over 90% in the poorest countries to 20% in upper-middle income countries. In Figure 3b, we observe that the negative IEC slope first increases in magnitude, between lower income to middle income countries, and then slightly decreases, between middle and upper-middle income countries. The average IEC slope is -10.2, implying a 1 percentage point reduction in informal budget share when household expenditure increases by 10%. Figure A3 shows that these patterns are similar when using the country-specific formality assignment rule by store type discussed in Section 3.2.

4.2 Differences in Informal Consumption across Households

Our micro database allows us to quantitatively investigate the main hypotheses proposed in the literature to explain differences in informal consumption between households. The first hypothesis is that poor and rich households differ in their characteristics which, through economies of scale and life-cycle patterns, impact where people shop (Deaton and Paxson (1998)). To measure how much of the IEC can be explained, we estimate the following regression in each country:

$$Share\ Informal_i = \beta * ln(expenditure_i) + \Gamma X_i + \varepsilon_i \tag{1}$$

where i indexes a household, X_i are household characteristics (household size and the age, education and gender of self-reported household head). Table 2 shows the

[13] Country level IECs are plotted for all 32 countries in the online appendix.

[14] Almås (2012) similarly finds a stable log linear relationship between food budget shares and household income around the world. For more disaggregated goods, however, Engel curves can be non-linear and vary across countries (Atkin et al., 2018a).

average of the slope coefficients β across countries. Relative to the specification without controls (column 1), accounting for household characteristics explains almost none of the IEC (column 2).

The second hypothesis is that poor households' access to formal stores is limited (Lagakos, 2016). To test this, we include controls for household location either with an indicator for rural area (column 3) or with survey block fixed effects (column 4).[15] We find that access matters, but only to a certain extent: controlling for rural location (block location) reduces the average slope by 16% (28%).

The third hypothesis is that of non-homothetic preferences: richer households spend more on goods predominantly sold in formal stores. To test this, columns 5 to 8 show product-level versions of (1) at increasingly narrow product levels.[16] Preferences across goods play an important role: controlling for food versus non-food lowers the slopes by 42% (column 5), and controlling for the 12 goods categories at COICOP 2-digit level accounts for 50% of the variation (column 6). Controlling for narrower goods categories only slightly reduces the slope further.

Column 9 in Table 2 combines all three hypothesis, which collectively account for 54% of the variation. Nonetheless, even with these extensive observable controls, the average IEC slope is -4.6 and remains statistically significant in all but three countries. The fourth hypothesis is that, within location and product categories, richer households value higher quality varieties which are more likely to be sold in formal stores. Such taste-based preferences are unobservable in the context of equation (1). Instead, we leverage the fact that in six countries of our sample, households are asked to report the main reason for choosing a place of purchase for each item. Table A1 indicates that households shop at informal stores for lower prices (column 1) and at formal stores for higher quality (column 2). This result holds within households, where formal (informal) purchases are more often motivated by higher quality (lower prices). We find that in each of the six countries, richer households are up to four times more likely to report quality as the main reason (see the online appendix).[17]

[15] Survey blocks are the most granular locations and contain on average 74 households in our surveys. The median survey block is representative on average of 52,900 people.

[16] Formally we estimate: $Share\,Informal_{ig} = \beta * ln(expenditure_g) + \alpha_g + \Gamma X_i + \varepsilon_{ig}$ where $Share\,Informal_{ig}$ is the share of household i's informal expenditure on good g and α_g are goods fixed effects. Observations are weighted by household weights and goods' expenditure shares.

[17] This is consistent with studies showing richer households spend more on branded goods in the United States (Faber and Fally, 2017) and on high-quality goods in Mexico (Atkin et al., 2018b).

This quality-price trade-off hypothesis implies that formal varieties of a good should be more expensive than informal varieties, reflecting quality differences. In the 21 countries where data permits it, we study the price difference in formal and informal stores within the most narrow good classification and location. We limit ourselves to food products to mitigate comparability issues and because food is often tax exempt. We estimate the formal price premium in each country:

$$ln(unit\ value)_{igmu} = \beta\ Formal_{igmu} + \mu_{gmu} + \epsilon_{igmu} \tag{2}$$

where $ln(unit\ value)_{igmu}$ is the unit value reported by household i, for good g, in location m, in units u, and $Formal_{igmu}$ equals one if the good is purchased in a formal store. μ_{gmu} are fixed effects at the good-location-unit level. On average, food prices are 6.7% higher in formal than informal stores (Table A2). This formal store premium is robust to excluding outliers and self-production, and to controlling for household characteristics. It is consistent with the hypothesis that formal stores sell high quality varieties at higher prices.[18]

This analysis suggests that non-homothetic preferences for quality and goods explain an important part of the IECs' downward slope, with some role for access. The relevance of these results for policy design depends on the cost for governments to observe the determinants of households' choices (type of goods, product quality). The type of good is relatively easy to observe: indeed, governments often set lower rates on food to relieve poorer households; however, we will see in Section 5 that the strong association between food and informal consumption drastically reduces the potential of such policies. On the contrary, product quality is costly to observe and can't be used directly in tax policy design.

4.3 Consumption Patterns of Food and Non-Food Formal Goods

To make consumption taxes more equitable, most countries set reduced rates or fully exempt food.[19] These policies are motivated by the steep downward slope of the food Engel curve, a pattern extensively documented together with its near

[18] We understate the true price-premium if other characteristics specific to formal stores reduce prices, such as productivity.

[19] Some countries apply reduced rates to all food goods and others target 'basic' food. For illustrative purpose, we follow the former approach. Targeting narrower items can improve redistribution, but increases the possibility for misreporting.

log linearity (Anker et al., 2011; Almås, 2012). It is thus relevant to investigate how the well-established food Engel curve changes once we focus on formal food consumption. Figure 4 shows for all countries in our sample the aggregate budget shares and Engel curves slopes, for total food consumption (Figures 4a and 4b), and for formal food consumption (Figures 4c and 4d). While the total budget share spent on food decreases as countries get richer, the budget share on formal food *increases*. Within country, the food Engel curves' slopes are strongly negative, while the formal food Engel curves have small *positive* slopes in poor and middle income countries, and become negative in upper-middle income countries.

Figures 4e-4f show the aggregate budget shares and Engel curve slopes for formal non-food. The budget share devoted to formal non-food consumption strongly grows across countries, from less than 20% in the lowest-income countries to 60% in upper-middle income countries. Similarly, the positive formal non-food Engel curve slopes triples over development.

5 How Progressive are Consumption Taxes?

In this section we analyze how these novel consumption patterns determine the progressivity of consumption taxes in the average country and across countries.

5.1 Progressivity in the Average Developing Country

Intuition A tax policy is progressive if the effective tax rate (ratio of taxes paid to household income) increases with household income. Following the literature on income tagging (Akerlof, 1978), we focus on the correlation between the budget share spent on a good and household income. The larger this correlation (in absolute terms), the better the consumption of that good is at tagging income. Thus, taxing a good whose consumption is positively correlated with income or exempting a negatively correlated good are both progressive policies. To build intuition, consider a good with an Engel curve that is upward sloping and linear with respect to log household income (for example formal goods and non-food goods). The progressivity achieved by taxing this good increases with the slope of its Engel curve and decreases with its aggregate budget share. Holding the aggregate budget share constant, an increase in the steepness of the Engel curve slope makes the good a better tag of income. Similarly, holding the (positive)

slope constant, a decrease in the aggregate budget share makes the good a better income tag, since it is more likely that a purchase of that good is made by a rich household. Thus, taxing the good achieves more progressivity.

Set-up We study the progressivity of three tax policy scenarios. Scenario #1 applies a uniform tax rate on all goods consumed from formal retailers to illustrate the progressivity of our new informality channel. Scenario #2 sets a zero tax rate on food and only taxes formal non-food consumption. This captures the combined progressivity impact of the (de-facto) exemption of informal stores and the policy exemption of food products. Comparing scenario #2 to #1 reveals the marginal progressivity gain from exempting food, when only formal consumption can be taxed. Scenario #3 applies a zero rate on food goods, but assumes that taxes are paid on expenditures from all store types, including home production. This corresponds to the unrealistic assumption of perfect enforcement that has implicitly been the focus of prior studies in developing countries. Comparing the progressivity achieved under scenario #3 to that achieved when moving from #1 to #2 captures how much failing to account for informality leads to incorrect conclusions about the redistributive potential of food exemptions. For each scenario, we assume that the government sets rates to collect 10% of total consumption in taxes. This maintains revenue collected constant across scenarios.[20]

Results Figure 5 shows, for the three scenarios, the effective tax rates faced by households in each decile of the total expenditure distribution, on average across countries. We obtain three main results. First, taxing only formal consumption makes consumption taxes progressive. Under scenario #1, the effective tax rate sharply increases across deciles: the richest quintile pays twice as much taxes (as a share of income) as the poorest quintile. This is because the informality Engel curves are downward sloping in all countries (Figure 3b). Second, the marginal progressivity achieved by exempting food when only formal consumption is taxable is limited: when moving to scenario #2 from scenario #1, the increase in progressivity is quantitatively small. This is because the formal food share doesn't always decrease with income (contrarily to overall food share): its Engel curve

[20]The simulations are mechanical: households consumption behavior is not affected by the level of tax rates. We assume no household savings; this is relaxed in Section 5.3.

slope is positive in most countries (Figure 4f). Third, the progressivity gain from exempting food in the realistic setting with informal consumption is much smaller than the progressivity gain in unrealistic scenario #3 with perfect enforcement. Indeed, a naive policymaker who does not account for informal consumption would overestimate the progressivity gains by a factor of 2.6: we can see this by comparing the ratio of top to bottom quintile effective tax rates under scenario #3 to the difference in ratio between scenario #2 and scenario #1. This over-estimation arises because food Engel curves are strongly negative everywhere (Figure 4b) while formal food Engel curves are either positive or mildly negative (Figure 4d).

5.2 Progressivity Across Countries

Measuring tagging potential We now consider how the progressivity of taxing different goods changes across countries. Rather than displaying effective tax rates across income distributions in all countries, we create an aggregate measure of a good's income tagging potential as the log of the ratio of the budget share spent on that good by households in the richest quintile relative to the poorest quintile. The log transformation implies that a positive (negative) value corresponds to a progressive (regressive) tax base, and a value of zero is distributionally neutral (i.e. budget shares of rich and poor are equal).[21]

Cross-country differences Figure 6 plots the tagging potential of different goods against the countries' income per capita. Figure 6a shows that taxing formal consumption is progressive in all 32 countries (log ratio above zero) and the progressivity is markedly higher in low-income countries than middle-income countries. This result is driven by two counteracting forces: the slope of the IEC grows in absolute value across development which increases progressivity, but this 'slope effect' is dominated by the 'base effect' whereby the large reduction in informal budget share over development decreases progressivity (Figures 3a-3b).

Figures 6b and 6c study the progressivity of taxing formal food and formal non-food. The results for formal food are intriguing (Figure 6b): in the poorest countries, formal food is a *progressive* tax base; its progressivity falls over devel-

[21] This transformation also ensures that the income tagging potential of goods' consumption is symmetric around zero: a good with a budget share ratio of $1/2$ has the same tagging potential (in absolute magnitude) as a good with a ratio of 2.

opment, and taxing formal food only becomes regressive in upper middle-income countries. This surprising result is driven by two factors. The first is the change in sign of the formal food Engel curve slope, which goes from small positive values in the poorest countries—because richer households consume most formal food products—to negative values in upper-middle income countries. The second is the increase in the formal food budget share across development (Figures 4c-4d). In contrast, formal non-food is positively correlated with income in all countries, though the budget share increases with country income (Figures 4e-4f). This makes taxing formal non-food a progressive policy everywhere, but its progressivity slightly declines over development (Figure 6c).

These patterns imply that taxing just formal food or just formal non-food are both *progressive* policies in the poorest countries: Figure 6d shows that formal food (dotted orange line) is just as strong a tag for rich households as formal non-food (solid orange line). Consequently, exempting food from taxes benefits rich households in these countries and worsens progressivity. Over development, the tagging potential gap widens between formal non-food and formal food consumption, but exempting food from taxation only starts to produce progressivity gains in upper middle-income countries. These pronounced differences across development explain why the progressivity gain from exempting food is limited in the average developing country (Figure 5).

Comparison with unrealistic perfect enforcement scenario To illustrate the importance of accounting for informal consumption, Figures 6e and 6f plot the progressivity of taxing *all* food and non-food goods, respectively. This corresponds to the unrealistic setting with perfect tax enforcement in both modern and traditional stores. In this setting, taxing food is regressive everywhere, and increasingly so over development as the food budget share decreases while the food Engel curve slope remains constant (Figures 4a-4b). Symmetrically, taxing non-food is progressive everywhere. These patterns are consistent with the general prior in the literature, but in sharp contrast with the nuanced results obtained in the realistic setting with limited enforcement in traditional stores (Figures 6b-6c).

Finally, we contrast the progressivity gains from exempting food in the realistic versus unrealistic scenarios (Figure 6d). The progressivity achieved by food exemption can be measured as the difference between the tagging potential of

non-food (solid lines) and food (dotted lines). In turn, we assess the difference between the unrealistic and realistic scenarios in the tagging potential difference between food and non-food. This 'difference-in-differences' between the two scenarios captures how the progressivity gains from exempting food is overestimated when making the unrealistic assumption of perfect enforcement. We find that overestimation occurs in all countries; it is largest in the poorest countries, and reduces over development. The existence of informal consumption thus dampens the progressivity gains from food exemption in all developing countries, but the overestimation error decreases as more consumption shifts into modern retailing. It is in the lowest-income countries that a naive policy evaluation of food exemption would lead to particularly erroneous conclusions.

5.3 Robustness

Our baseline assignment is country-invariant and assumes that all modern (traditional) stores are formal (informal). We relax this assumption by using instead the country-specific formality shares of each store type, that we estimate using the WBEIS-Euromonitor data. The main results are robust to this alternative assignment, which only leads to a small decrease in progressivity both in the average country (Table A3) and across countries (Figure A4).

Our baseline results use total expenditure to proxy for income, thus implicitly assuming that households do not save. Savings decrease effective consumption tax rates and reduce progressivity if the savings rate increases with income.[22] Measuring savings in expenditure surveys is challenging in developing countries where income is poorly measured (Deaton, 1997). Instead, we use the Global Findex Database (Demirguc-Kunt et al., 2018) to measure the share of households that save by income decile. We then apply a homogeneous savings rate (conditional on saving) to match aggregate savings in the economy: this produces a country-specific distribution of savings across the income distribution. Allowing for distributional savings slightly decreases progressivity in all scenarios, as expected, but the main results remain unchanged (Table A3 and Figure A4).

[22] Annual income overstates consumption taxes' regressivity: consumption depends on lifetime income which is less volatile than annual income (Poterba, 1989).

6 Implications for Optimal Tax Policy

This section studies the implications of the novel consumption patterns for tax policy design. We extend the Ramsey model of commodity taxation (Diamond, 1975) to a context in which informal varieties cannot be taxed. This allows us to derive simple formulae for optimal tax rates which we then calibrate to our data.

6.1 Model

Households There is a continuum of mass 1 of households i with heterogeneous exogenous incomes y^i. Households have preferences over j goods, and for each good over two varieties v: $v = 0$ indicates a variety produced in the informal sector, which cannot be taxed, $v = 1$ a variety produced in the formal sector. Producer prices q_{jv} are exogenous. Consistent with our main formality assignment, consumer prices are given by $p_{j1} = q_{j1}(1 + t_j)$, where t_j is the tax on good j, and $p_{j0} = q_{j0}$. We write $v(p, y^i)$ the indirect utility of household i, s^i_{jv} the budget share spent by household i on variety v of good j, $s^i_j = s^i_{j0} + s^i_{j1}$ the budget share spent on good j, and ϵ_j the price elasticity of demand for good j.

We assume that formal and informal varieties are substitutes. This introduces an additional efficiency cost of taxation compared to a model in which all varieties can be taxed: as the price of the formal variety rises, households substitute to informal varieties which leads to a tax revenue loss (see Appendix D for details). We further assume for convenience that demand elasticities are equal across households and that elasticities of substitution across goods that are taxed differently are equal to zero. This assumption is reasonable given that we only allow for differentiated tax rates across large product categories (e.g. food versus non-food), but it is relaxed in Appendix D for completeness.[23]

Government preferences The government chooses the tax rates t_j to maximize:

$$W = \int_i G(v(p, y^i))di + \mu \sum_j t_j q_{j1} x_{j1} \tag{3}$$

where $x_{j1} = \int_i x^i_{j1}(p, y^i)$ is total consumption of the formal variety of good

[23] We study rate differentiation across 12 large product categories in Section 7.2; the elasticity of substitution is likely to remain small in this setting (e.g. between health products and clothing).

j. Government preferences are characterized by μ, the marginal value of public funds, and $G()$, an increasing and concave social welfare function. 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 $\bar{g}$ its average (see Saez and Santcheva, 2016). We assume g^i falls with income, and $\mu = \bar{g}$. The latter is a convenience assumption corresponding to a government that taxes only if it enables redistribution; we relax it in Appendix Section D.2.

Optimal uniform commodity taxation Consider a uniform tax on all goods, $t_j = t$ for all j. Writing $\tau = \frac{t}{1+t}$, welfare maximization yields:

$$\tau^* = \frac{\int_i (\bar{g} - g^i)\phi^i \frac{s_1^i}{s_1} di}{-\epsilon_1 \bar{g}} \tag{4}$$

where $s_1 = \sum_j \int_i s_{j1}^i di$ is the aggregate budget share spent on all formal varieties, $\phi^i = \frac{y^i}{\bar{y}}$ is the ratio of household i's income relative to average income $\bar{y}$ and ϵ_1 is the uncompensated price elasticity of demand for all formal varieties. Equation (4) shows that the optimal uniform rate is increasing in the co-variance between household income and the formal budget share: the more richer households spend on formal varieties relative to the poor, the more redistribution is obtained from taxing only formal varieties. The existence of an informal sector therefore increases the optimal uniform rate, absent efficiency considerations: downward-sloping IECs indicate that the correlation between total formal consumption and income is higher than that between total consumption and income. In other words, more redistribution is achieved from taxing only formal varieties than from taxing all varieties uniformly.

The optimal rate decreases in the absolute value of the uncompensated price elasticity of demand for formal varieties: the more households respond to changes in formal prices by consuming fewer formal varieties, the higher the efficiency cost of taxing only formal varieties. Appendix D shows that this effect is increasing in the elasticity of substitution in consumption across varieties and the share of the informal variety in total consumption of the product. The more households are willing to substitute to informal varieties, the more demand for the formal variety responds to an increase in the tax and the higher the efficiency cost of taxing only formal varieties. This implies that efficiency considerations will push the optimal

rate on formal varieties down relative to the optimal rate on all varieties, and more so the larger the informal retail sector.

Optimal rate differentiation Consider now a government that sets different rates on goods. The optimal rate on good j is:

$$\tau_j^* = \frac{\int_i (\bar{g} - g^i)\phi^i \frac{s_{j1}^i}{s_{j1}} di}{-\epsilon_{j1}\bar{g}} \tag{5}$$

This expression shows that the optimal rate on good j is increasing in the covariance between household income and budget share spent on the formal variety of good j. It is decreasing in the absolute value of the uncompensated price elasticity of the formal variety, which itself is increasing with the informal budget share for the good and the elasticity of substitution between varieties.

6.2 Calibrated Optimal Tax Policy

Table 3 summarizes our choices of parameters when calibrating the optimal tax rates defined in equations (4) and (5). We use our data to calibrate the budget shares for each good, variety and country, and the slopes of the Engel curves to calibrate income elasticities.[24] A key parameter is the cross-variety compensated price elasticity which governs the substitution between formal and informal varieties: we consider a range of [1,2] in line with estimates in Faber and Fally (2017) and Atkin et al. (2018b); we use 1.5 as our baseline value for all goods, but also present results obtained with a value of zero (implying no substitution from formal to informal varieties) to illustrate the implications of ignoring the efficiency costs due to informal retailers. We set a value of -0.7 for the own-price compensated elasticity of goods. Together, these parameters yield values for the own-price uncompensated elasticity of goods in the $[-2.2, -0.7]$ range, consistent with the literature (Deaton et al., 1994). Finally, we calibrate the government's social welfare weights such that the optimal uniform rate is on average 18% across

[24]The slopes of the Engel curves identify the income elasticities under the assumption that the observed correlation between income and budget shares is causal. To approximate a causal income effect, we estimate the slopes of the Engel curves using a specification with the full set of controls used in Table 2 except for product fixed effects. Our income elasticities are therefore not confounded by changes in demand due to household characteristics (demographics, location, etc) that are correlated with income. See the notes to Table 3 for more details.

the countries in our sample.[25] This matches the statutory consumption tax rates often set in practice (see Appendix D.2 for details). Demand parameters and government preferences are the same in all countries, so that any cross-country variation in optimal rates is driven by differences in consumption patterns.[26]

Figure 7 presents calibrated rates against countries' income per capita. The first panel plots optimal uniform rates. We see that optimal rates are lower in the poorest countries (15%) than in middle-income countries (20%). This result arises because the efficiency gains due to the shrinking share of formal consumption (Figure 3) push optimal rates up; this effect is stronger than the falling progressivity of taxing formal consumption (Figure 6), which pushes optimal rates down. Indeed, Figure A6 shows that the uniform rate is steeply decreasing across higher levels of income per capita if we assume that informal consumption has no efficiency costs (through zero cross-variety substitution). For a plausible range of elasticity of substitution values, however, uniform rates are non-decreasing with income per capita (Figure A6).

The second panel of Figure 7 plots the ratio of the optimal food to non-food rates. A lower value indicates a higher optimal subsidy on food; a value above one indicates that the optimal policy taxes food more heavily than non-food. This panel reveals that the optimal rate on food is 20% lower than on non-food in low-income countries but 40% lower in middle-income countries. This effect is driven by the progressivity patterns across goods (Figure 6): once informal consumption is accounted for, the tagging potential of exempting food is limited in poorer countries but higher in middle-income countries so the optimal policy subsidizes food less in poorer countries. In some of the poorest countries, Figure 7 shows that food should not be subsidized relative to non-food (ratio above one): this is because the benefits from a food subsidy would accrue disproportionately to richer households who consume the bulk of formal food. These results hold for

[25] Our baseline results assume, in line with our model, that the value of public funds is equal to the average social welfare weight ($\mu = \bar{g}$). We relax this assumption in Figure A8, discussed in Appendix Section D.2. Increasing μ increases optimal rates, as expected, but patterns with respect to economic development are unchanged.

[26] Income inequality can also play a role: optimal rates will be larger in settings where inequality is higher. We gauge the importance of inequality for the cross-country patterns in Figure A5, which plots the optimal uniform rates under the naive assumption that all varieties can be taxed. These rates only vary because of differences in inequality: contrary to our main result in Figure 7, we see no systematic cross-country pattern. See also Appendix D.2.

all plausible values of the cross-variety elasticity of substitution (Figure A7).

7 Impacts on Inequality and Extensions

In this section we quantify the inequality effects of optimal consumption taxes. We then extend our model and calibration to consider rate differentiation across multiple goods and the presence of a direct personal income tax.

7.1 Inequality Reduction from Optimal Consumption Taxes

Our inequality metric is the percent change in Gini from the pre-tax to the net-of-tax expenditure distribution.[27] In the case of a uniform tax rate, the inequality impact depends on the level of the optimal rate, the progressivity of taxing formal consumption, and the share of formal consumption. For differentiated rates, the inequality impact depends, in addition, on the progressivity gains from taxing formal non-food versus food differently, and on the size of the respective bases.

Figure 8 shows for each country the percent change in Gini from applying optimal uniform taxes (Figure 8a) and optimal food and non-food rates (Figure 8b). Two key results emerge. First, the inequality reduction achieved due to informal consumption is sizeable, at 1.92% on average. Redistributive gains increase across countries, from 1% on average in the poorest countries to 3% in upper-middle income countries. This is due to both higher tax rates and higher formal budget shares in richer countries, and despite the falling progressivity of taxing formal consumption as countries become richer. Second, the marginal redistribution gain from rate-differentiation is limited: on average, the Gini effect increases from 1.9% to 2.6%. Thus, in the average country, the redistribution potential of consumption taxes is primarily due to distributional differences in informal consumption rather than policy choices (rate-differentiation). However, the marginal inequality reduction from rate-differentiation grows with per capita income, thanks to both the increased progressivity of reduced rates on formal food (Figure 6) and more rate differentiation (Figure 7).

We gauge the magnitude of our results by comparing them to Commitment to Equity (CEQ) studies in 25 developing countries (Lustig, 2018). Importantly, CEQ evaluates actual rather than optimal policies and usually does not account

[27] We only consider the direct effect of the tax system and do not reallocate revenue collected.

for informal consumption. Notwithstanding, CEQ findings reflect the consensus view on the limited redistributive role of consumption taxes. Indeed, the average inequality reduction from these taxes in CEQ is 0.6%; over three times smaller than our average estimates (1.92-2.64%). In fact, our results are more comparable to the inequality reduction achieved by *actual* income taxes and social security in CEQ (2.6% on average). We discuss the role of direct taxes in Section 7.3.

Robustness Table A4 reports robustness checks. Panel (a) shows the average Gini reduction and Panel (b) the ratio of Gini reduction in middle-income countries over that in low-income ones. First, we vary the cross-variety price elasticity of demand: a lower value yields more inequality reduction as governments can set higher taxes without inducing much substitution towards informal varieties. This mainly benefits low-income countries with larger informal sectors. Second, we present results obtained using our estimated country-specific formality shares by store type instead of our baseline formality assignments, and allowing for country-specific distributional savings rates. These checks reduce the average inequality impacts by 3-10% across scenarios but leave our main results unchanged.

7.2 Further Rate Differentiation

We have focused on inequality achieved by a policy that differentiates rates optimally across two types of products (food versus non-food), a fair approximation of tax policies in place in developing countries. To study how further rate differentiation impacts inequality, we calibrate optimal tax rates for each of the twelve main goods categories of the UN COICOP classification (food, clothing, etc.), for each country. Figure A9 shows that the dispersion in optimal rates increases as countries get richer, mirroring the pattern with only two rates. The average inequality reduction achieved by rate differentiation across twelve goods is 3.2%, which is 20% higher than with two goods (Table A4). However, further rate differentiation achieves no additional inequality reduction in low-income countries (Figure A9). Since further rate differentiation induces tax evasion and administrative costs (Ebrill and Keen, 2001), this exercise cautions against extensively deploying rate differentiation for equity motives – particularly in low-income countries.

7.3 Interaction with Direct Taxes on Income

A central result in public finance is that redistribution is better achieved through direct rather than indirect taxes (Atkinson and Stiglitz, 1976). However, this result assumes that income taxes are perfectly enforceable, which is at odds with developing countries' reality (Jensen, 2019). Indirect tax instruments serve a redistributive role as soon as the possibility of evading income taxes is taken into account (Huang and Rios, 2016).

We incorporate a personal income tax (PIT) in our model in Appendix D.3, characterized by an exemption threshold (above which people pay income taxes) and a single marginal tax rate. The PIT lowers disposable income above the threshold which affects optimal consumption tax policy in two ways. It decreases the welfare gains from taxing richer households via taxes on goods with steep Engel curves (as these households are already taxed by the PIT). It also lowers rich households' consumption of these goods, which decreases their tagging potential. Both effects reduce optimal commodity taxes, and more so when the exemption threshold is lower. Jensen (2019) shows that the PIT exemption threshold gradually declines as income per capita increases, leading to an expansion of the PIT base, but the top marginal tax rate remains fairly constant. In our calibration, we use the data from Jensen (2019) to predict the size of the PIT base for our sample of countries. We assign to all countries the top marginal rate observed in our data (50%). This choice is conservative and overstates the redistribution achieved by PIT, thereby understating the redistributive potential of consumption taxes.[28]

Figure A10 shows that the presence of the PIT lowers the uniform consumption tax rates from 18% to 16% on average. These lower rates reduce the inequality impacts: accounting for PIT implies that the average inequality reduction from consumption taxes falls from 1.9% to 1.6% with a uniform rate, and from 2.6% to 2.1% with differentiated rates (Table A4). These results are obtained in a setting that overstates *actual* PIT policies.[29] A richer setting with optimal PIT policies lies beyond the scope of this paper, but we note that an optimal PIT is unlikely to be

[28] We model an ideal PIT where all households above the threshold fully comply with the (top) marginal rate. This further understates the redistributive potential of consumption taxes. Indeed, our simulations show a larger Gini reduction from the modelled PIT system —5.5% on average.

[29] Calibration results obtained using each country's actual top marginal PIT rate yield very similar conclusions: the optimal uniform rate is 16.8%; the average inequality reduction falls to 1.7% under a uniform rate and to 2.3% under differentiated rates.

substantially more redistributive than the constrained PIT we calibrate, thus leaving our qualitative results unchanged, for two reasons. First, optimal marginal rates are unlikely to be well above the 50% rate used here. Second, optimal PIT threshold levels will likely be high in many developing countries, leading to a narrow tax base, due to the enforceability constraints discussed in Jensen (2019).

8 Incidence Considerations

Our main formality assignment is based on the stylized assumption of 100% pass-through of taxes in modern stores and 0% pass-through in traditional stores. In this section, we provide additional evidence on this assumption and discuss how our results change under more complex incidence assumptions.

8.1 Pass-through by Store Type: Evidence from Mexico's VAT reform

Set-up and results A concern with our baseline assignment is that factors other than the tax status (formal or informal) of a store may impact the extent of tax pass-through to consumers. To gauge the importance of this concern, we directly estimate the VAT pass-through in modern and traditional stores using product-level price data and a VAT reform in Mexico. In January 2014, the VAT increased from 11% to 16% in border areas to equal the non-border rate which remained at 16%. Our identification strategy compares prices in each store type (modern, traditional) between border and non-border areas over time. The difference-in-differences design recovers the causal impact of the VAT increase on retail prices in a store type if there are no changes to unobservable factors which coincide with the timing of the reform and differentially impact prices in border versus non-border stores. We implement a flexible regression model which includes all month dummies between January 2013 and January 2015. The omitted period is the reform announcement date (August 2013), which allows us to inspect anticipation effects ahead of the implementation in January 2014. Formally, we estimate

$$lnp_{glt} = \sum_t \beta_t^s (Border_l * Period_t) + \mu_g + \mu_l + \mu_t + \epsilon_{glt} \tag{6}$$

where lnp_{glt} is the log of the VAT-inclusive price of product g in location l at time t; μ_g, μ_l, and μ_t are fixed effects at the product, location and time level,

respectively. $Border_l = 1$ if the location is in the border areas. Standard errors are clustered at the product-location level. The coefficient of interest is β_t^s, which we estimate separately in the two store type samples $s = (Traditional, Modern)$.

To avoid potential cross-location spillovers, we focus on non-tradable products following the classification in Mariscal and Werner (2018), and in line with recent studies on VAT pass-through (Benzarti and Carloni, 2019; Harju et al., 2018).[30] Our estimation sample contains 291,840 (160,368) product-location-month price observations in traditional (modern) stores. Appendix C.2 contains more details.

Figure 9 plots the results. In modern stores, prices evolve similarly in treatment and control stores up to the implementation month. We observe a sharp price increase in the immediate post implementation period in treated stores. When averaging $\widehat{\beta_t^M}$ over all months, the modern-store pass-through $\widehat{\beta^M}$ is 77% (3.85 percentage points). In contrast, the price impact in traditional stores is muted: prices modestly increase in the immediate post implementation month. Beyond this impact, the post-implementation trend reverts to the small and steadily declining pre-trend. For this reason, measuring the pass-through as the average over all post-implementation months would create bias (Rambachan and Roth, 2021). Instead, we measure the pass-through by comparing coefficients between the immediate post-implementation month and the last pre-implementation month, giving a traditional store pass-through $\widehat{\beta^T}$ of 14% (0.70 percentage points).[31]

We consider the robustness of our results by assigning the pass-through rates estimated in Mexico to modern and traditional stores in all our countries. Results are shown in the last column of Table A3 (for progressivity) and Table A4 (for inequality). Compared to our baseline, this alternative assignment assumes a less than full pass-through in modern stores, which unambiguously dampens the redistributive effects of consumption taxes.[32] However, it also assumes a higher pass-through in traditional stores than in the baseline setting, which has an am-

[30] Because the treatment varies at the local level, including tradables would lead us to underestimate the level of VAT pass-through if households buy tradables across local areas. Combined with the exclusion of non-taxable goods, the focus on taxable non-tradable goods implies that the set of products we focus on are predominantly services.

[31] This captures the average effect in traditional stores if the full price impact occurs in the immediate implementation month; the absence of dynamic price-impacts in modern stores suggests this assumption is plausible. Estimates are robust to winsorizing the data; extending the sample to more periods; and, restricting control areas to be geographically close to border areas.

[32] A lower pass-through in modern stores decreases the overall progressivity of consumption taxes and the size of the tax base, which lowers the average effective rates.

biguous effect on inequality reduction: on one hand, it decreases the progressivity of a consumption tax; on the other hand, it increases the size of the tax base which magnifies the redistributive effect of any progressive tax. The negative effects dominate overall, leading to a roughly 30% decrease in the inequality effect.

Discussion Estimating pass-through by store type for every country is beyond the scope of this paper: it requires detailed consumer price data by store type and VAT policy reforms that provide credible research designs in 32 countries.[33] Here we discuss how pass-through rates may differ in countries other than Mexico.

For informal stores, the theoretical literature highlights that, under a VAT, informal retailers may purchase some of their inputs from VAT-registered suppliers and pass on this tax-cost to final consumer prices (Keen, 2008). In Appendix D.4, we formalize this mechanism in a setting where formal and informal retailers engage in monopolistic competition. The model predicts that the pass-through in informal stores equals the share of intermediate products purchased from formal suppliers in total costs.[34] This share is likely to be small for several reasons. First, a VAT system disincentivizes trade between VAT and non-VAT firms, which leads to segmentation between formal and informal supply chains (De Paula and Scheinkman, 2010; Gadenne et al., 2020). Studies find limited evidence of links with upstream formal firms: the Informal Economy Monitoring Study surveys informal retailers in 7 countries and finds that only 20.5% report purchases form formal suppliers (Mahadevia et al., 2014). For informal firms in 6 West African cities, Bohme and Thiele (2014) similarly estimate that only 8% source any input from formal firms. Second, home-production represents 36% of consumption in traditional stores in the average country in our sample (Figure A2). Since households are unlikely to use formal inputs to produce at home, the pass-through of taxes is likely close to zero for this large category of traditional consumption.

Recent studies, albeit all in high-income countries, find that the VAT pass-through to (formal) prices may not be full and depends on firm and market characteristics (Harju et al., 2018; Benzarti et al., 2020). When the pass-through

[33]The set-up would require: (1) time-series data on VAT-inclusive prices; (2) store type classification; (3) a reform to the VAT rate at a treatment-level which minimizes spillovers. These requirements are stringent; our search suggests that (2) is a binding constraint in many countries.

[34]This share is 10% when measured in the Mexican retail census. The 10% model-predicted pass-through is close to the estimated pass-through (14%) in Mexican traditional stores (Fig.9).

to prices is incomplete, the tax burden is shared among consumers, firm owners and workers (Benzarti and Carloni, 2019). In Mexico, the incomplete pass-through of 77% in modern stores implies that some of the VAT burden could have been passed on to employees (via lower wages) or to store owners (via reduced profits).[35] Our qualitative results hold in this richer incidence setting, as long as formal store owners and workers are richer than the average household. Using the surveys' employment modules for 19 of our sample countries, we indeed find that formal retail workers are on average 50% richer than the median household.[36]

More generally, our results hold qualitatively under any positive difference in pass-through between modern and traditional stores, though the quantitative results depend on the precise pass-through levels. As future research measures pass-through by store type in more countries, these estimates can be combined with our model and publicly available data on informal consumption to refine the optimal rates and quantitative inequality impacts in each developing country.

8.2 Other considerations

We focus on consumption taxes, but retailers may be liable for other taxes. Informal retailers sometimes pay market fees and presumptive taxes, while formal retailers pay taxes on corporate income and property. Accounting for such additional taxes likely leads to a higher overall effective tax rate for formal stores than for informal stores which, if passed on to prices, further increases tax progressivity. Entry and exit of firms in the modern and traditional sectors may also affect the final tax incidence. A VAT rate increase may cause some formal retailers to become informal, putting downward pressure on informal prices.

Finally, the societal desirability of a traditional retail sector depends on multiple factors not considered here, including impacts on health and employment. Traditional retailers can improve food access for poor households due to their numerous locations and smaller bundles (Crush, 2018); but storage limitations in these stores also cause food hazards and low nutrient quality (Hopkins, 2006). Moreover, traditional retail employs 15-20% of the workforce in many developing

[35]Some modern stores in Mexico may not remit VAT in practice, which would also lead to an incomplete pass-through when estimated in the sample of all modern stores.

[36]The income-difference between formally employed workers versus the median worker is most pronounced in low-income countries (Figure A11). We measure formality in the surveys as any health or pension contributions by the household head.

countries, but modern retailers pay higher wages (Cardiff-Hicks et al., 2014). The lack of consensus is reflected in the ongoing debate over regulating modern retail entry; country policies vary from full liberalization (Latin America) to prohibitive barriers (Southeast Asia). Our study contributes to this debate by investigating how informal retailers affect the equity of tax systems.

9 Conclusion

In this paper, we harmonize expenditure surveys from 32 developing countries which contain the store type for each transaction. We assign store types to the informal or formal sector using a robust assignment rule, and measure the informal budget share at the household level. We find that informal budget shares steeply fall with household income in every country. Contrary to the consensus, consumption taxes are progressive in developing countries and optimal commodity taxes lower inequality by 2-3%, as much as actual personal income taxes. Our results have sharp implications for the use of reduced rates on necessities, a widespread policy around the world. We show that differentiating rates across goods has limited redistributive potential once informal consumption is accounted for, particularly in low-income countries.

Tax administrations recognize how taxing it is to tax small firms. As a result, they often focus enforcement on large firms (Basri et al., 2019) and exempt firms below a threshold (Keen and Mintz, 2004). Going forward, the growth of digital technologies may lower enforcement and compliance costs and make it possible to bring smaller firms into the tax net (Gupta et al., 2017). Our results do not imply that efforts to tax small firms should be abandoned, but caution that the benefits from reducing the size of the informal sector should be weighed against equity costs. Policy decisions—such as the location of the exemption threshold—should consider distributional impacts in addition to compliance costs.

References

Akerlof, G. A. (1978): "The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning," *American Economic Review*, 68, 8–19.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019): "Regressive Sin Taxes,

with an Application to the Optimal Soda Tax," *Quarterly Journal of Economics*, 134, 1557–1626.

Almås, I. (2012): "International Income Inequality: Measuring PPP Bias by Estimating Engel Curves for Food," *American Economic Review*, 102, 1093–1117.

Almunia, M., L. Liu, B. Lockwood, and E. Tam (2019): "VAT Notches, Voluntary Registration, and Bunching," *Review of Economics and Statistics*.

Alvaredo, F. and L. Gasparini (2015): *Recent Trends in Inequality and Poverty in Developing Countries*, vol. 2 of *Handbook of Income Distribution*, Elsevier.

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," *NBER Working Paper*, No. 26890.

Atkin, D., B. Faber, and M. Gonzalez-Navarro (2018b): "Retail Globalization and Household Welfare," *Journal of Political Economy*, 126, 1–73.

Atkinson, A. and J. Stiglitz (1976): "The Design of Tax Structure: Direct versus Indirect Taxation," *Journal of Public Economics*, 6.

Basri, M. C., M. Felix, R. Hanna, and B. A. Olken (2019): "Tax Administration vs. Tax Rates," *NBER WP*, No. 26150.

Benzarti, Y. and D. Carloni (2019): "Who Really Benefits from Consumption Tax Cuts? Evidence from VAT Reform in France," *AEJ: Economic Policy*, 11.

Benzarti, Y., D. Carloni, J. Harju, and T. Kosonen (2020): "What Goes Up May Not Come Down: Asymmetric Incidence of Value-Added Taxes," *Journal of Political Economy*, 128, 4438–4474.

Bergeron, A., G. Tourek, and J. Weigel (2021): "The State Capacity Ceiling on Tax Rates: Evidence from Randomized Tax Abatements in the DRC," *WP*.

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.

Bohme, M. and R. Thiele (2014): "Informal-Formal Linkages and Informal Enterprise Performance in Urban West Africa," 25, 473–80.

Bronnenberg, B. J. and P. B. Ellickson (2015): "Adolescence and the Path to Maturity in Global Retail," *Journal of Economic Perspectives*, 29, 113–34.

Cardiff-Hicks, B., F. Lafontaine, and K. Shaw (2014): "Do Large Modern Retailers Pay Premium Wages?" *NBER Working Paper Series*.

Casaburi, L. and U. Troiano (2015): " Ghost-House Busters: Electoral Responses to an Anti–Tax Evasion Program *," *Quarterly Journal of Economics*, 131, 273–314.

Crush, J. (2018): *The 'Supermarketization' of Food Supply and Retail*, Routledge.

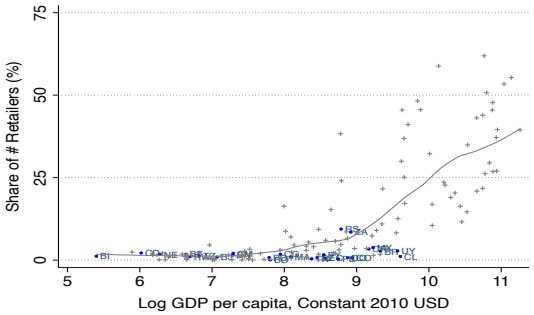
De Paula, A. and J. A. Scheinkman (2010): "Value-Added Taxes, Chain Effects, and Informality," *American Economic Journal: Macroeconomics*, 2, 195–221.

De Simone, R. (2020): "Tax Enforcement as Corporate Governance," *WP, LBS*.
Deaton, A. (1997): *The Analysis of Household Surveys*, The World Bank.
Deaton, A., K. Parikh, and S. Subramanian (1994): "Food Demand Pattern and Pricing Policy in Maharashtra," *Sarvekshana*, 17, 11–34.
Deaton, A. and C. Paxson (1998): "Economies of Scale, Household Size, and the Demand for Food," *Journal of Political Economy*, 106, 897–930.
Demirguc-Kunt, A., L. Klapper, D. Singer, S. Ansar, and J. Hess (2018): *The Global Findex Database 2017*, The World Bank.
Diamond, P. A. (1975): "A Many-Person Ramsey Tax Rule," *Journal of Public Economics*, 4, 335–342.
Ebrill, L. P. and M. Keen (2001): *Modern VAT*, International Monetary Fund.
Evenson, R., P. P. e. (2007): *Handbook of Agricultural Economics*, 3, North-Holland.
Faber, B. and T. Fally (2017): "Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data," *NBER Working Paper*, No. 23101.
Gadenne, L. (2017): "Tax Me, but Spend Wisely? Sources of Public Finance and Government Accountability," *American Economic Journal: Applied Economics*, 9, 274–314.
Gadenne, L., T. Nandi, and R. Rathelot (2020): "Taxation and Supplier Networks: Evidence from India," *CEPR Discussion Paper*, No. DP13971.
Gemmell, N. and O. Morrissey (2005): "Distribution and Poverty Impacts of Tax Reform in Developing Countries," *Development Policy Review*, 23, 131–144.
Gerard, F. and G. Gonzaga (2016): "Informal Labor and the Efficiency Cost of Social Programs," *NBER Working Paper*, No. 22608.
Gupta, S., M. Keen, A. Shah, and G. Verdier, eds. (2017): *Digital Revolutions in Public Finance*, International Monetary Fund.
Harju, J., T. Kosonen, and O. N. Skans (2018): "Firm types, price-setting strategies, and consumption-tax incidence," *Journal of Public Economics*, 165, 48–72.
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," *IFS Working Paper*, W18/11.
Hopkins, R., T. R. (2006): "The Supermarket Revolution in Developing Countries: Policies to Address Emerging Tensions Among Supermarkets, Suppliers, and Traditional Retailers," *European Journal of Development Research*, 18, 522–545.
Huang, J. and J. Rios (2016): "Optimal Tax Mix with Income Tax Non-Compliance," *Journal of Public Economics*, 144, 52–63.
Humphrey, J. (2007): "The Supermarket Revolution in Developing Countries," *Journal of Economic Geography*, 7, 433–450.
Jaravel, X. (2019): "The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector," *The Quarterly Journal of Economics*, 134, 715–783.
Jensen, A. (2019): "Employment Structure and the Rise of the Modern Tax system," *NBER Working Paper*, No. 25502.
Keen, M. (2008): "VAT, tariffs, and withholding: Border taxes and informality in

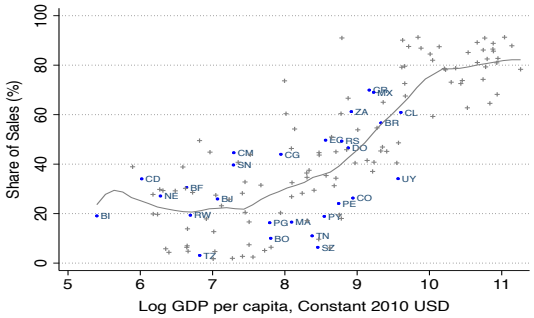
developing countries," *Journal of Public Economics*, 92, 1892–1906.
Keen, M. and J. Mintz (2004): "The Optimal Threshold for a Value-Added Tax," *Journal of Public Economics*, 88, 559 – 576.
Kleven, H. and M. Waseem (2013): "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *Quarterly Journal of Economics*, 128, 669–723.
Kleven, H. J. (2014): "How Can Scandinavians Tax So Much?" *Journal of Economic Perspectives*, 28, 77–98.
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.
Londono-Velez, J. and J. Avila-Mahecha (2021): "Enforcing Wealth Taxes in the Developing World: Quasi-Experimental Evidence from Colombia," *American Economic Review: Insights*, 3, 100–128.
Lustig, N. (2018): *Commitment to Equity Handbook: Estimating the Impact of Fiscal Policy on Inequality and Poverty*, Brookings Institution Press.
Mahadevia, D., S. Vyas, and A. Mishra (2014): "Street Vendors in Ahmedabad, India," *Informal Economic Monitoring Study Country Series*.
Mariscal, R. and A. Werner (2018): "The Price and Welfare Effects of the Value-Added Tax: Evidence from Mexico," *IMF Working Papers*, 18.
Naritomi, J. (2019): "Consumers as Tax Auditors," *American Economic Review*, 109.
Pomeranz, D. (2015): "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax," *American Economic Review*, 105, 2539–69.
Poterba, J. (1989): "Lifetime Incidence and the Distributional Burden of Excise Taxes," *American Economic Review, Papers and Proceedings*, 79.
Rambachan, A. and J. Roth (2021): "An Honest Approach to Parallel Trends," *Working Paper*.
Reardon, T., C.Timmer, and J. Berdegue (2003): "The Rise of Supermarkets in Africa and Asia," *American Journal of Agricultural Economics*, 85, 1140–1146.
Saez, E. and S. Santcheva (2016): "Generalized Social Marginal Welfare Weights for Optimal Tax Theory," *American Economic Review*, 106, 24–45.
Sah, R. K. (1983): "How Much Redistribution is Possible through Commodity Taxes?" *Journal of Public Economics*, 20, 89–101.
Ulyssea, G. (2018): "Firms, Informality, and Development: Theory and Evidence from Brazil," *American Economic Review*, 108, 2015–47.
Waseem, M. (2020): "Overclaimed Refunds, Undeclared Sales, and Invoice Mills," *CESifo Working Paper*, No. 8231.
Weigel, J. (2019): "The Participation Dividend of Taxation: How Citizens in Congo Engage with the State When It Tries to Tax Them," *Quarterly Journal of Economics*, 135.

Figure 1: Characteristics of Modern and Traditional Stores across Countries

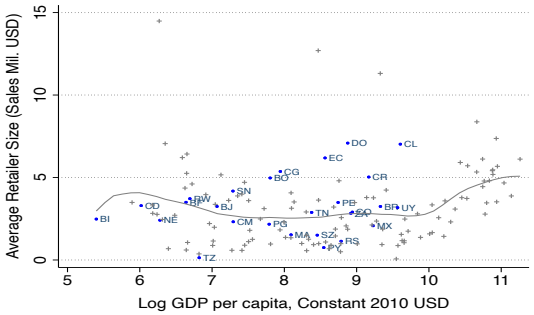
(a) Modern Stores, Share of Retailers



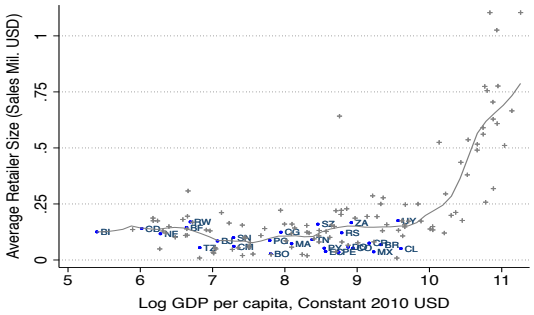
(b) Modern Stores, Market Share



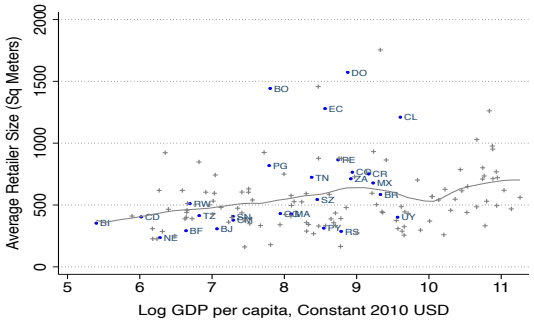
(c) Modern Stores, Average Sales



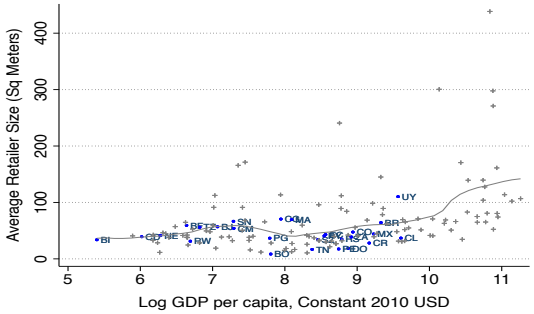
(d) Traditional Stores, Average Sales



(e) Modern Stores, Average Space

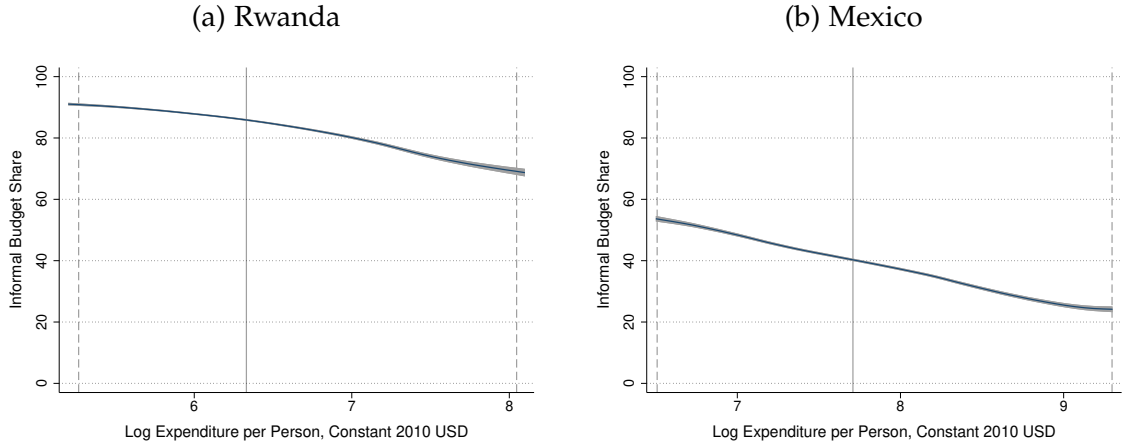


(f) Traditional Stores, Average Space



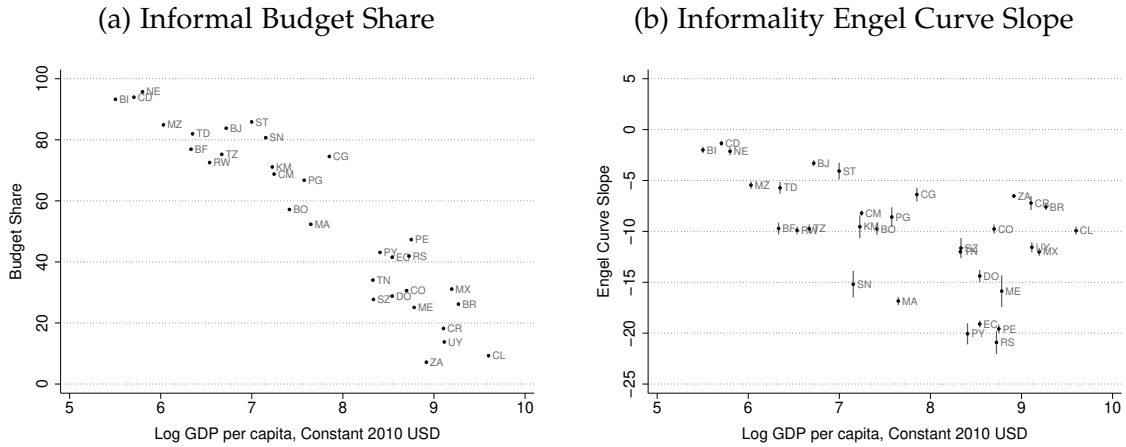
Notes: These panels plot characteristics of modern and traditional stores in 2014 across 189 countries at different levels of log GDP per capita. All figures are based on retail reports from Euromonitor International. Traditional stores include non-brick and mortar stores and small corner and convenience stores. Modern stores include specialized stores and large stores. Panels a-b plot the modern-store share in total number of retailers and total retail sales. Panels c-d plot the average sales size for modern and traditional stores, measured in USD millions. Panels e-f plot the average floor space of modern and traditional stores, measured in square meters. The blue dots denote the 32 countries in our data, where we observe store type for each transaction. More details in Section 3.1.

Figure 2: Selected Informality Engel Curves



Notes: These panels show the local polynomial fit of the Informality Engel Curve (IEC) in Rwanda and Mexico, constructed from household level data. The informal budget share is on the vertical axis. Log per person total expenditure is on the horizontal axis. The shaded area around the polynomial fit corresponds to the 95% confidence interval. The solid vertical line denotes the median of each country's expenditure distribution, while the dotted lines correspond to the 5th and 95th percentiles. More details in Section 4.1. See the online appendix for each country's IEC.

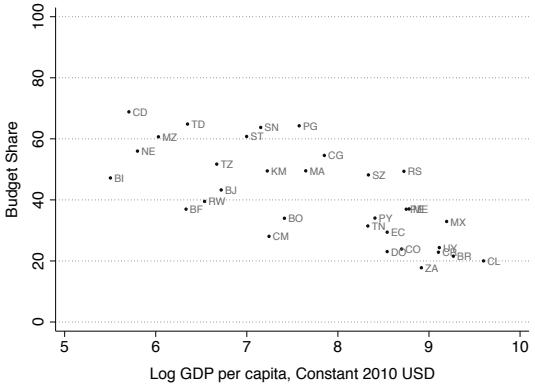
Figure 3: Informal Expenditure Across Countries



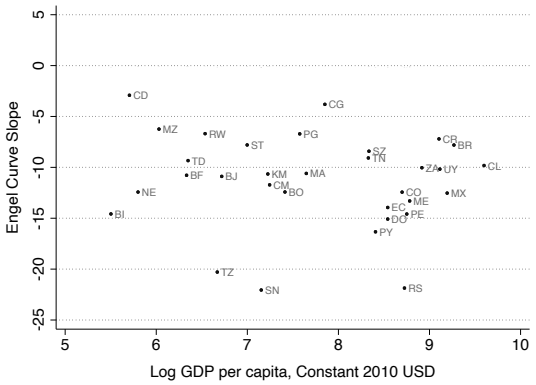
Notes: Figure 3a plots the aggregate informal budget share against log GDP per capita for each country. Figure 3b shows the slope of the informality Engel curves against log per capita GDP. The bars denote the 95% confidence interval of the slope coefficient. More details in Section 4.1.

Figure 4: Expenditure on Different Goods Across Countries

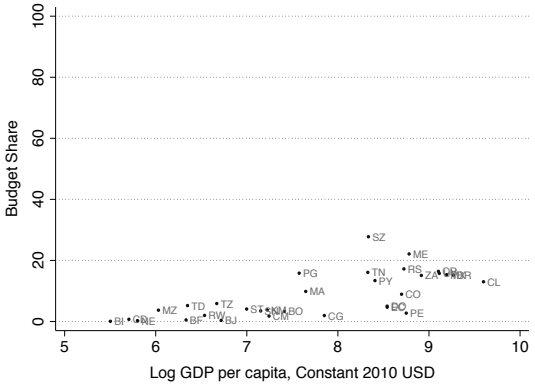
(a) Food Budget Share



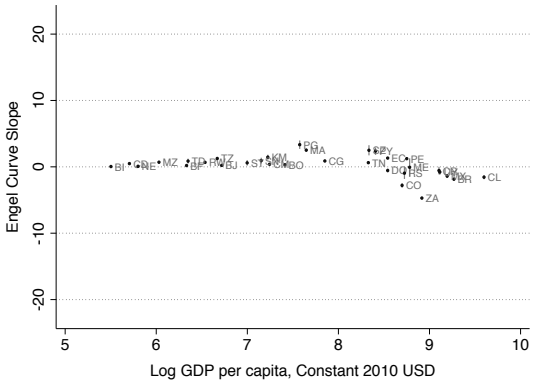
(b) Food Engel Curve Slope



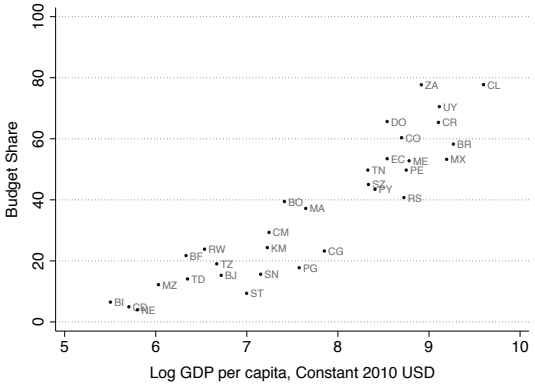
(c) Formal Food Budget Share



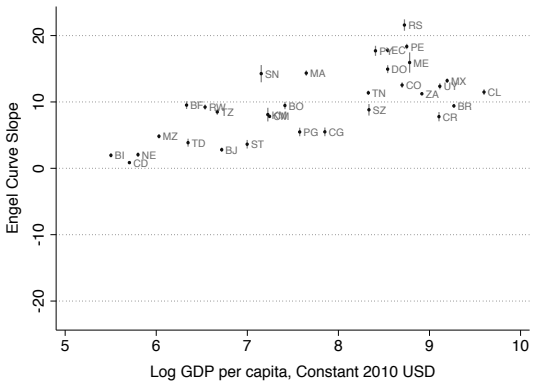
(d) Formal Food Engel Curve Slope



(e) Formal Non-Food Budget Share

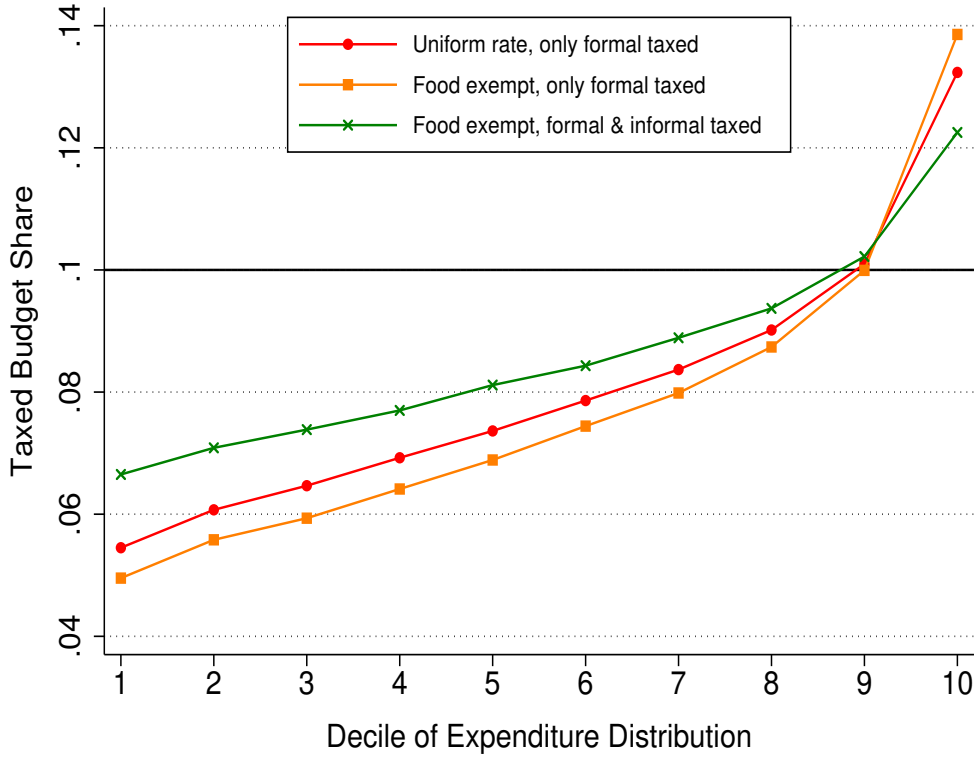


(f) Formal Non-Food Engel Curve Slope



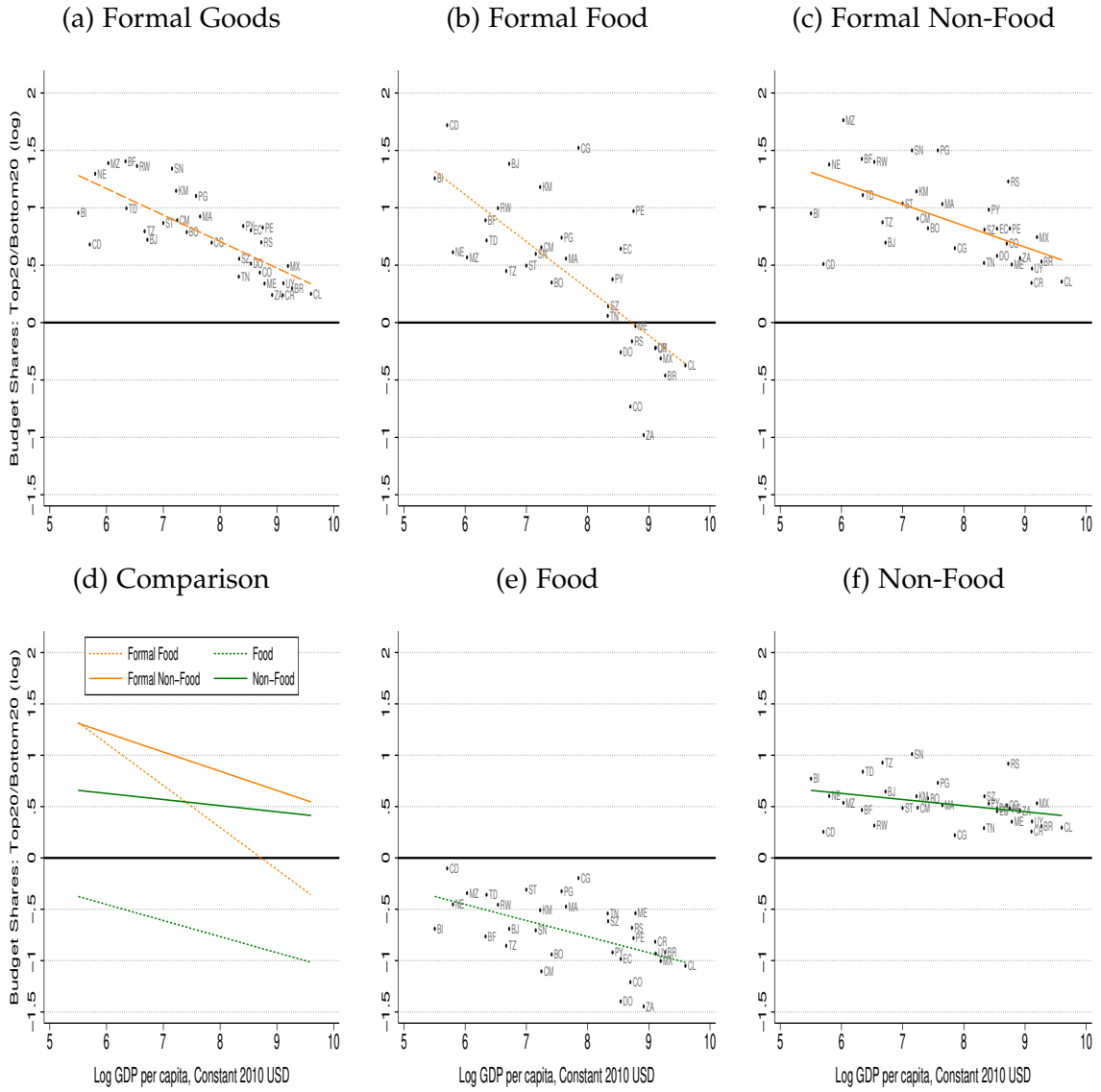
Notes: This figure shows the aggregate budgets shares (left panels) and Engel curves slopes (right panels) against log GDP per capita for three types of consumption goods: food (panels a-b), formal food (panels c-d), and formal non-food (panels e-f). More details in Section 4.3.

Figure 5: Progressivity of Tax Policy Scenarios in the Average Country



Notes: This figure plots the share of expenditures paid in taxes (effective tax rates) by expenditure decile, for the three tax policy scenarios described in Section 5.1. The three scenarios are simulated in all 32 countries and each point corresponds to the average effective tax rate of each decile across countries. Each scenario imposes that the government collects 10% of total consumption in taxes and assumes that households do not respond to taxes by changing their consumption choices (mechanical simulations). The black horizontal line at 10% is thus the effective tax rate when all consumption —formal and informal— is taxable at a uniform rate. The red circled line corresponds to a scenario where a uniform tax is levied on all goods consumed from formal retailers. The orange squared line corresponds to a scenario where food is zero-rated and only formal non-food consumption is taxed. The difference between the orange squared line and the red circled line captures the marginal progressivity gains of tax exempting food when informal consumption is accounted for. The green crossed line corresponds to the scenario with a zero rate on food goods, but with taxes paid on consumption in both formal and informal stores. The difference between the green crossed line and the horizontal black line captures the marginal progressivity gain of exempting food in the unrealistic scenario with perfect tax enforcement in all stores.

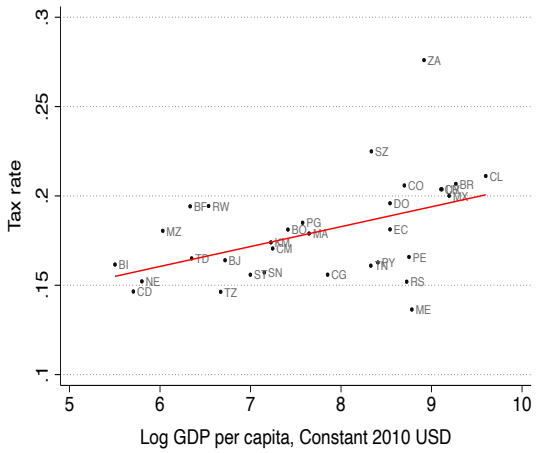
Figure 6: Progressivity of Taxing Different Goods across Countries



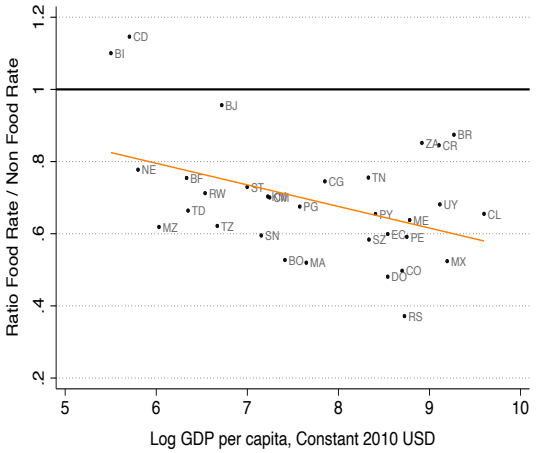
Notes: This figure plots the log of the ratio of the budget shares spent on a good by the richest quintile of households relative to that of the poorest quintile, against the country's log per capita GDP. The log transformation implies that a positive value of the budget share ratio corresponds to a progressive tax base, a negative value to a regressive one and a zero value to a neutral tax base (i.e. the budget share of rich and poor are equal). In each panel, the slope corresponds to the best linear fit across countries. Each panel considers a different tax base: (a) formal consumption, (b) formal food, (c) formal non-food, (e) food (formal and informal), (f) non-food (formal and informal). Panel (d) shows the linear fit lines of panels (b) and (c) in orange, and panels (e) and (f) in green, to directly compare the progressivity of taxing different bases. The dashed (solid) lines correspond to the setting where only formal goods (both informal and formal goods) are taxed. Thus, in panel (d) the difference between the dashed and solid line (holding color constant) measures the progressivity gains due to informal consumption, while the difference between colored lines (holding line patterns constant) measures the progressivity gains from exempting food. More details in Section 5.2.

Figure 7: Optimal Tax Rates Across Countries

(a) Optimal Uniform Rates



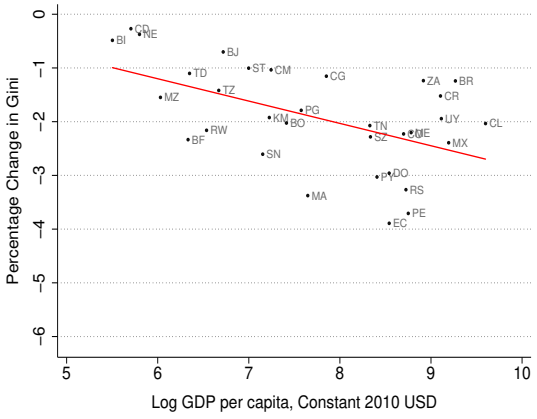
(b) Optimal Rate Differentiation



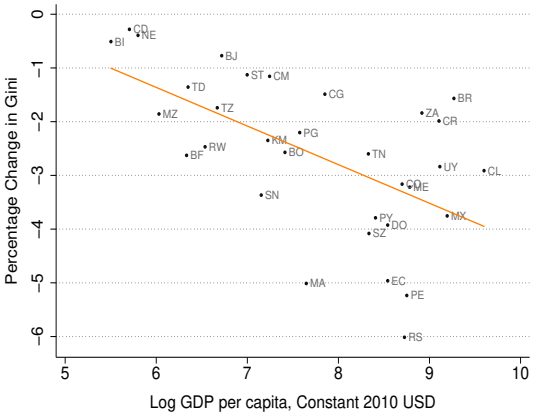
Notes: Panel (a) plots the optimal uniform tax rate as a function of a country's log GDP per capita, calibrated based on equation (4). Panel (b) plots the optimal level of rate differentiation between food and non-food products (measured as the ratio of optimal rate on food to optimal rate on non-food) as a function of a country's log GDP per capita. The optimal food and non-food rates are calibrated based on equation (5). All calibration parameters take the baseline values specified in Table 3. The lines correspond to the best linear fits.

Figure 8: Inequality Reduction from Optimal Tax Policy across Countries

(a) Uniform Rates



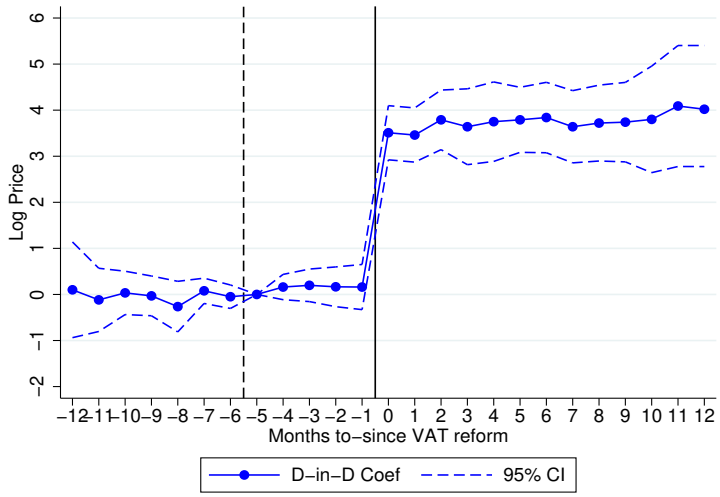
(b) Rate Differentiation



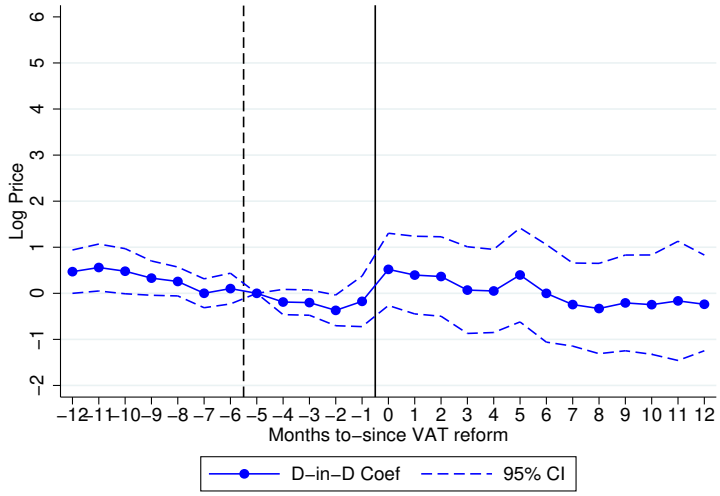
Notes: The figures plot the percentage change in the Gini coefficient from applying optimal commodity tax rates which account for informal consumption, against a country's log GDP per capita. The Gini coefficients are measured using percentiles of the pre-tax and post-tax expenditure distributions. Panel (a) corresponds to the uniform tax rate scenario and panel (b) corresponds to the scenario with differentiated rates for food and non-food goods. Lines denote the best linear fit. More details in Section 7.1.

Figure 9: Pass-through by store type: Evidence from Mexican VAT Reform

(a) Pass-through in Modern Stores



(b) Pass-through in Traditional Stores



Notes: This figure shows the pass-through of taxes to final consumer prices in modern stores (panel a) and in traditional stores (panel b), following a 5 percentage point increase in the VAT rate. Each panel plots the difference in differences regression coefficients β_t from estimating equation 6. The dashed lines correspond to the 95% confidence interval, where standard errors are clustered at the product-location level. The modern (traditional) sample has 160,368 (291,840) product-location-month price observations between January 2013 and January 2015. The vertical lines indicate the relevant dates: the dashed line denotes the date when the reform was announced (August 2013) and the solid line denotes the actual reform implementation date (January 2014), when the VAT rate increased from 11% to 16% in border areas. The omitted period is the announcement date, which allows us to inspect for anticipation effects in the intermediary period between announcement and reform implementation. Months are counted relative to the date of implementation of the reform (January 2014).

Table 1: Household Expenditure Surveys

Country	Code	Survey	Year	GDP pc (USD)	# Households	# Items/Hhld
Benin	BJ	EMICOV	2015	828	19871	32
Bolivia	BO	ECH	2004	1658	9149	49
Brazil	BR	POF	2009	10595	56025	41
Burkina Faso	BF	EICVM	2009	563	8404	152
Burundi	BI	ECVM	2014	245	6681	90
Cameroon	CM	ECAM	2014	1400	10303	81
Chad	TD	ECOSIT	2003	572	6697	94
Chile	CL	EPF	2017	14749	15239	129
Colombia	CO	ENIG	2007	5999	42373	60
Comoros	KM	EDMC	2013	1373	3131	82
Congo DRC	CD	E123	2005	301	12098	107
Congo Rep	CG	ECOM	2005	2569	5002	85
Costa Rica	CR	ENIGH	2014	8994	5705	68
Dominican Rep	DO	ENIGH	2007	5121	8363	88
Ecuador	EC	ENIGHUR	2012	5122	39617	89
Eswatini	SZ	HIES	2010	4169	3167	44
Mexico	MX	ENIGH	2014	9839	19479	61
Montenegro	ME	HBS	2009	6516	1223	149
Morocco	MA	ENCDM	2001	2095	14243	90
Mozambique	MZ	IOF	2009	416	10832	221
Niger	NE	ENCBM	2007	330	4000	192
Papua NG	PG	HIES	2010	1949	3810	111
Paraguay	PY	EIGCV	2011	4479	5417	88
Peru	PE	ENAHO	2017	6315	43545	78
Rwanda	RW	EICV	2014	690	14416	54
Sao Tome	ST	IOF	2010	1095	3545	100
Senegal	SN	EDMC	2008	1278	2503	299
Serbia	RS	HBS	2015	6155	6531	105
South Africa	ZA	IES	2011	7455	25328	44
Tanzania	TZ	HBS	2012	788	10186	318
Tunisia	TN	ENBCNV	2010	4142	11281	139
Uruguay	UY	ENIGH	2005	9079	7043	77

Notes: This table provides information on the surveys used in the 32 countries in our sample. Code refers to the country-code acronym which we use in figures. The original name of the survey is provided. GDP per capita is in PPP USD in the year of the survey, obtained from the World Bank Development Indicators. The sample size refers to the number of households in the survey, and the number of items is the number of expenditure items reported on average per household. More details in Section 2.1.

Table 2: Average Slopes of the Informality Engel Curves

Specification:	**Main**		**Geography**		**Product Codes**				**All**
Avg. of 32 Countries	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(Negative of) Slope	10.2	11.0	9.5	8.6	7.2	6.7	6.4	5.9	4.6
Confidence Interval	[9.6,10.8]	[10.3,11.6]	[8.9,10.2]	[7.8,9.3]	[6.6,7.8]	[6.2,7.1]	[5.9,6.9]	[5.4,6.3]	[4.1,5.0]
# of p-values < 0.05	32	32	32	32	32	31	31	31	29
R^2 adjusted	0.20	0.22	0.27	0.43	0.44	0.52	0.52	0.52	0.55
Household Characteristics		X	X	X	X	X	X	X	X
Urban/Rural			X						
Survey Blocks				X					X
Food Products					X				
COICOP 2-dig						X			
COICOP 3-dig							X		
COICOP 4-dig								X	X

Notes: This table shows the (negative of the) average slope of the Informality Engel Curves across countries for different specifications. Column 1 reports the slopes estimated from the following regression: $Share\ Informal_{ip} = \beta_0 + \beta_1 ln(expenditure_i) + \varepsilon_{ip}$ where $Share\ Informal_{ip}$ is the share of household i's informal expenditure on product p. Each observation is weighted using household survey weights and the expenditure share of the product. The average of lower and upper bound of 95% confidence intervals in brackets, calculates using robust standard errors. Column 2 augments this regression with controls for household characteristics (household size, age, gender, education of head). Column 3 (4), adds fixed effects for urban/rural (survey enumeration blocks). Column 5 instead adds fixed effects for food versus non-food products. Columns 6/7/8 instead add fixed effects for product codes at 2nd/3rd/4th level of the COICOP classification. Column 9 adds household characteristics and fixed effects for survey blocks and COICOP-4. More details in Section 4.2.

Table 3: Baseline Calibration Parameters

Parameter	Value	Source
Budget shares s_j^i and s_{j1}^i	Varying	Observed in our data
Household income (scaled) ϕ^i	Varying	Observed in our data
Income elasticities of goods η_j	Food: 0.7, Non-food: 1.2	Estimated from our data
Income elasticities of formal varieties η_{j1}	Food: 1.05, Non-food: 1.3, All goods: 1.3	Estimated from our data
Informal share of consumption α_j	Varying	Observed in our data
Cross-variety compensated elasticity	1.5	Faber and Fally (2017); Atkin et al. (2018b)[2]
Own-price compensated elasticity ϵ^C	-0.7	Deaton et al. (1994)[3]
Government preferences g^i	1-10	Average uniform tax rate of 18%[4]

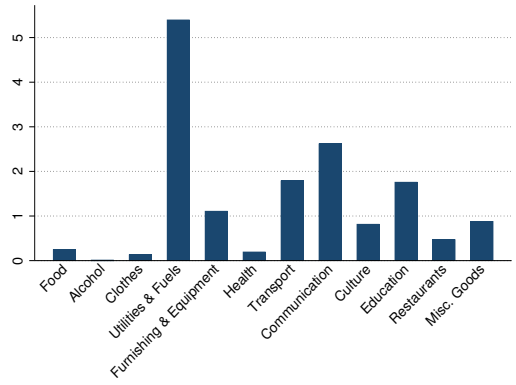
Notes: This table shows the parameters used to calibrate the optimal commodity tax rates in Section 6.2.
[1] For each product category k (food, non-food, formal food, formal non-food), we run the regression $s_{ki} = \beta_k ln(expenditure_i) + \Gamma X + \epsilon_i$, where X includes all the controls used in Column 5 of Table 2. We then obtain for each country an estimate of the income elasticity for k using $\eta_k = 1 + \hat{\beta}_k / s_k$, where s_k is the category expenditure share. We calibrate income effects using the average value across countries for each category.
[2] For the cross-variety price elasticity (parameter $\tilde{\epsilon}^C$ in Appendix Section D), we use estimates of the elasticity of substitution σ across store types in consumption obtained by Faber and Fally (2017); Atkin et al. (2018b) which are in the $[2, 4]$ range. With a CES utility function, we obtain $\tilde{\epsilon}^C = \sigma s_0$, where s_0 is the aggregate budget share spent in the informal sector (equal to 0.5 on average in our sample).
[3] Our choice of value for ϵ^C together with our estimated income elasticities and observed budget shares yield uncompensated own-price elasticities for goods in the $[-2, -0.5]$ range, in line with estimates obtained by Deaton et al. (1994) in developing countries.
[4] To match the average tax rate of 18% across countries, welfare weights fall for each decile in steps of approximately one. Thus, $g^i \approx 10$ for the first decile, $g^i \approx 9$ for the second decile, $g^i \approx 8$ for the third decile, ..., $g^i = 1$ for the tenth decile. Together with our other calibration choices, this yields an average optimal uniform rate of 18%, with country-specific rates in the 15-20% range, in line with observed consumption tax rates in developing countries.

Online Appendix
"Informality, Consumption Taxes and Redistribution"

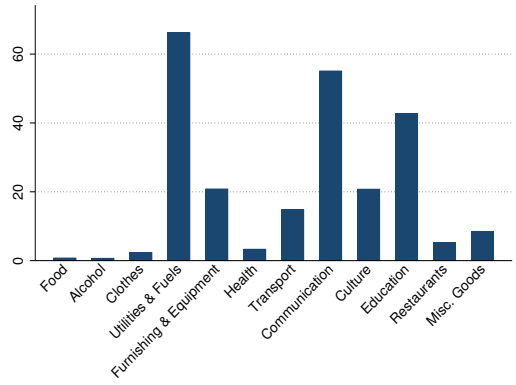
A Additional Figures and Tables

Figure A1: Unspecified Places of Purchase by Good

(a) Percentage of Total Expenditure



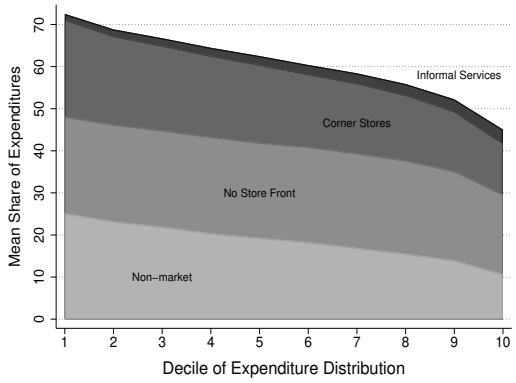
(b) Percentage of Goods' Expenditure



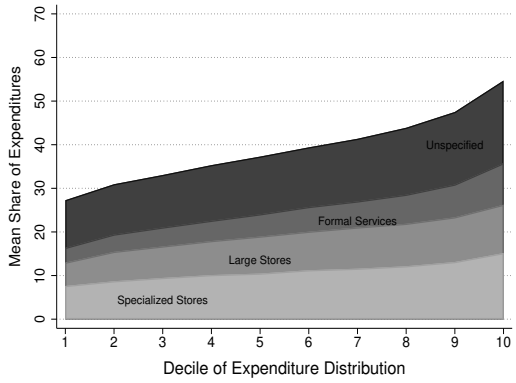
Notes: These panels show the share of expenditures with an unspecified place of purchase by good type (COICOP-2 digit) on average across the 32 countries of the sample, as discussed in Section 2.2.

Figure A2: Average Expenditure of Each Decile By Formality Assignment

(a) Informal Places of Purchase



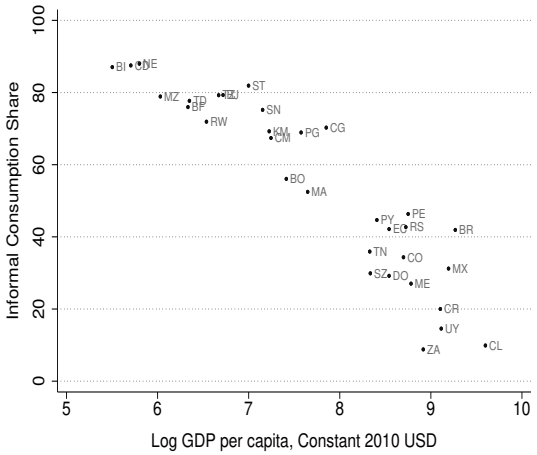
(b) Formal Places of Purchase



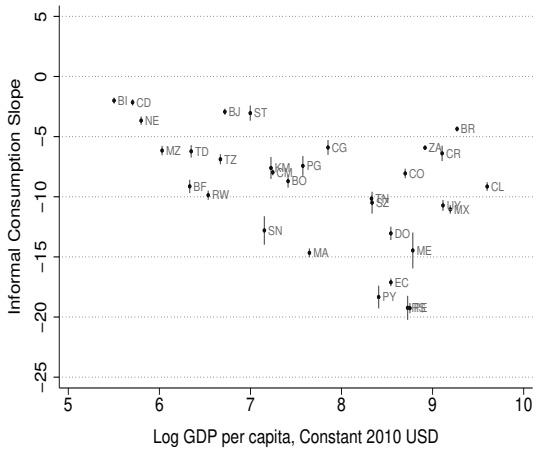
Notes: This figure shows the average cross-country budget-share by expenditure decile andd type of retailer, following the retailer taxonomy described in Section 3.2. Panel (a) shows the places of purchase classified as traditional and informal and Panel (b) shows the places of purchase classified as modern and formal in the central scenario of the paper.

Figure A3: Informality Engel Curves under Alternative Formality Assignment

(a) Budget Shares (Country Specific)

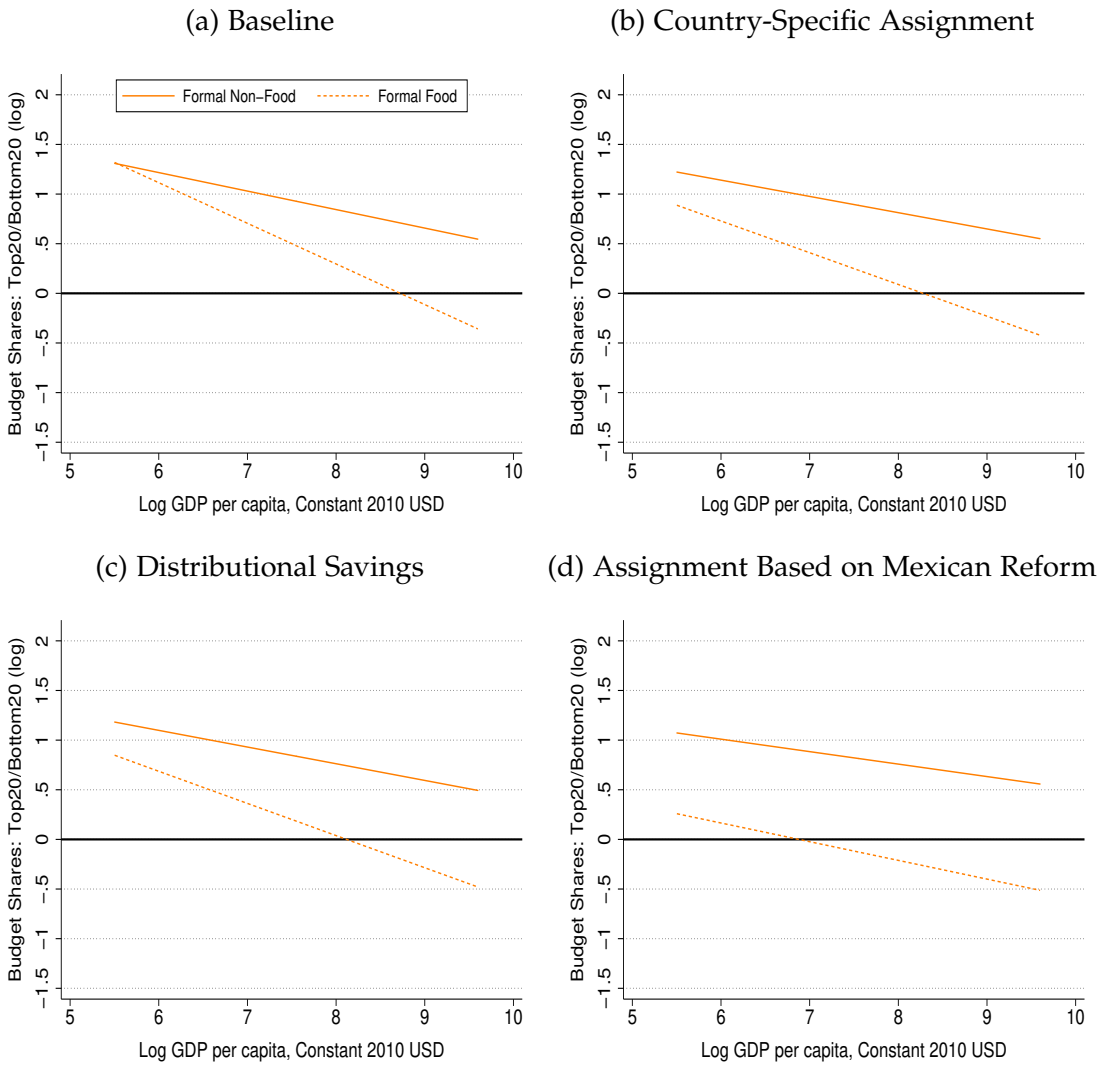


(b) Slopes (Country Specific)



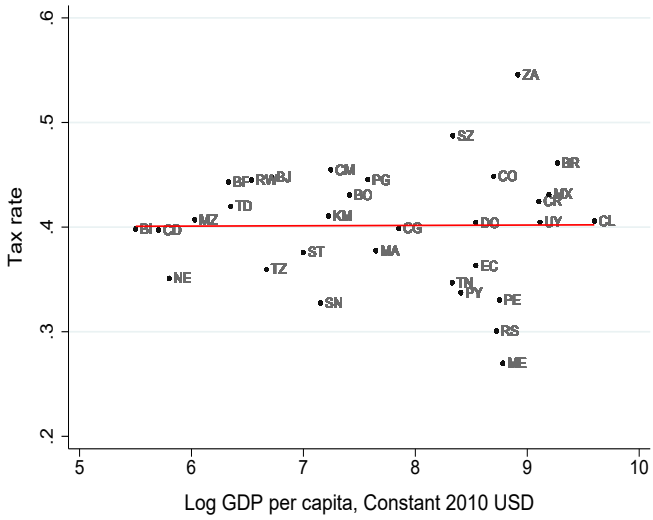
Notes: This figure is constructed similarly to Figure 3, but uses a country-specific assignment of store types to formality. The country-specific assignment differs from the baseline assignment by assigning a probability of formality to modern and traditional stores in each country, using data from the World Bank Enterprise Surveys and the Eurmonitor Retail reports. The graphs show informal consumption's aggregate budget shares (panel a) and Engel curves (panel b) against log GDP per capita. The bars correspond to the 95% confidence interval of the slope coefficient.

Figure A4: Progressivity of Tax Bases over Development in Different Scenarios



Notes: This figure shows the progressivity of taxing formal non-food goods and formal-food goods under the different assumptions discussed in Section 5.3. Panel (a) repeats the baseline scenario for comparison. Panel (b) uses the country specific probabilities of formality for modern and traditional stores. Panel (c) allows for country and decile specific distributional savings, such that the tax base is now total income rather than total expenditure. Panel (d) uses the tax pass-through estimates to modern and traditional stores obtained from the VAT reform in Mexico. In all panels, the vertical axis shows the ratio (in log) of the budget shares spent on a good by the richest quintile over that of the poorest quintile. The horizontal axis is log per capita GDP of the country. The log transformation implies that a positive value of the budget share ratio corresponds to a progressive tax base, a negative value to a regressive base and a zero value to a neutral tax base (i.e. budget shares of rich and poor are equal). We show the slopes which represent the best linear fit on the cross-country observations.

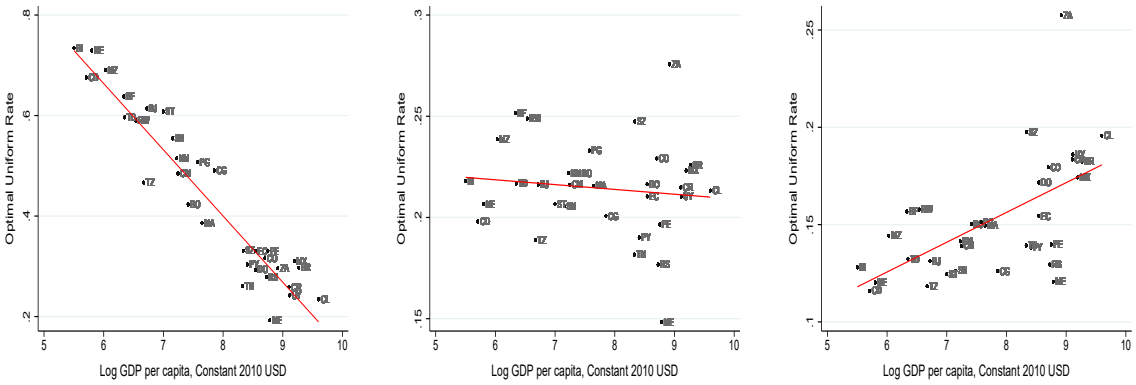
Figure A5: Optimal Uniform Rates with no Informal Sector



Notes: This figure plots the optimal uniform rates obtained under the assumption that all varieties of all goods are taxed equally, i.e. there is no informal sector, against the country's log GDP per capita. More details in Section 6.2.

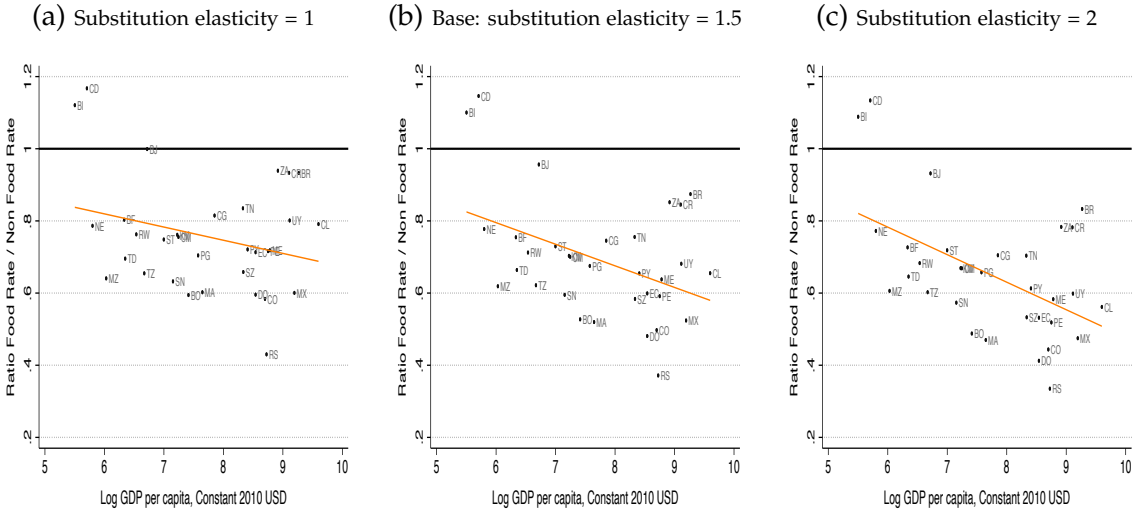
Figure A6: Optimal Uniform Rates: Varying Elasticity of Substitution

(a) Substitution elasticity = 0 (b) Substitution elasticity = 1 (c) Substitution elasticity = 2



Notes: These panels show the optimal uniform rates, calibrated using expression (4), for varying values of the elasticity of substitution in consumption between formal and informal varieties: 0 in the left-hand graph (meaning no efficiency cost of the informal sector); taxation); 1 in the middle graph; and, 2 in the right-hand graph. The lines correspond to the best fit across countries. More details in Section 6.2.

Figure A7: Optimal Rate Differentiation: Varying Elasticity of Substitution

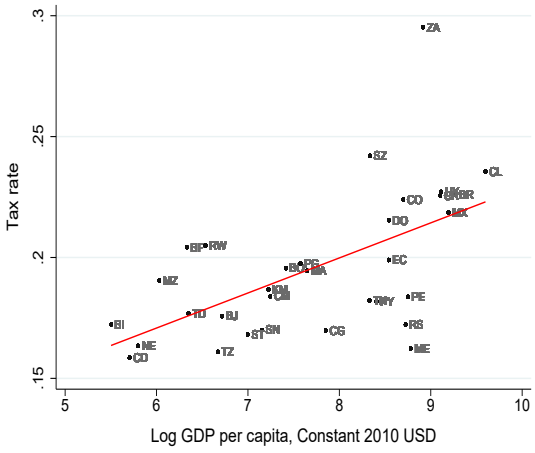


Notes: These panels plot the ratio of the calibrated optimal tax rate on food over that of non-food goods, when only formal varieties are taxed, against the country's log GDP per capita. Optimal rates are calibrated using expression (5). A ratio of 1 indicates that the optimal rates of food and non food are equal, and a ratio below 1 that it is optimal to set a lower rate on food goods. The cross-variety elasticity of substitution varies across graphs: 1 in the left-graph; 1.5 in the middle graph; 2 in the right-graph. The lines correspond to the best fit across countries. More details in Section 6.2.

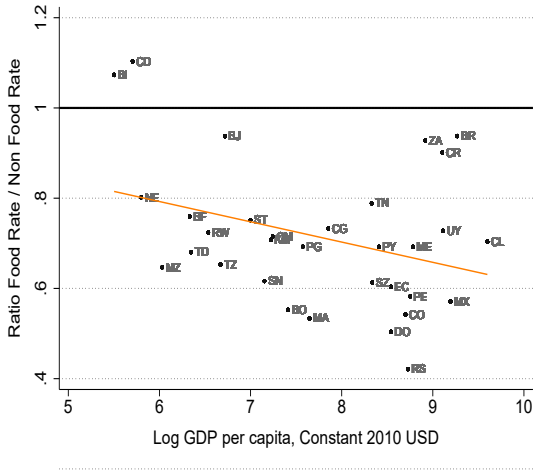
Figure A8: Optimal Tax Rates for Different Government Preferences

High μ

(a) Optimal Uniform Rate

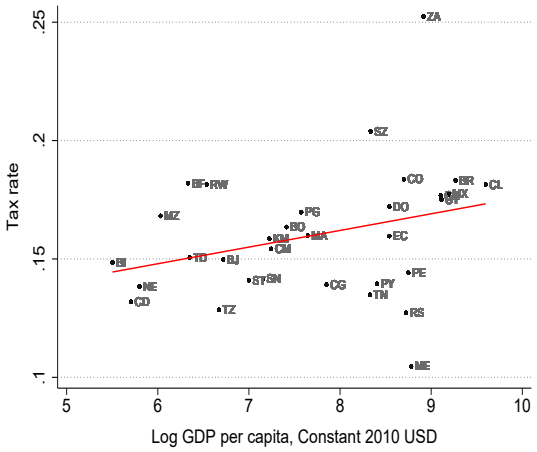


(b) Optimal Rate Differentiation

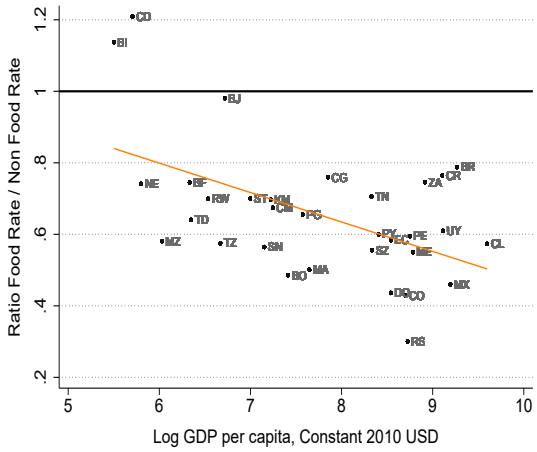


Low μ

(c) Optimal Uniform Rate



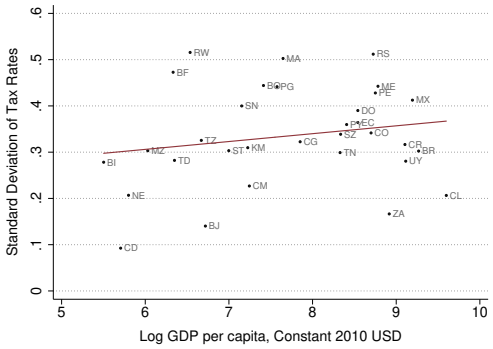
(d) Optimal Rate Differentiation



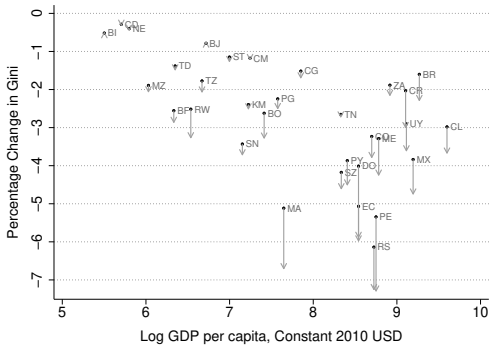
Notes: These panels plot optimal tax policies: optimal uniform rates (left panels) and optimal rate differentiation (ratio of optimal food rate to optimal non-food rate, right panels). The top two panels assume that $\mu = 1.1\bar{g}$ and the bottom two panels assume that $\mu = 0.9\bar{g}$. The lines correspond to the best linear fits across countries.

Figure A9: Tax Rate Differentiation Across Twelve Product Categories

(a) Tax Rates Dispersion



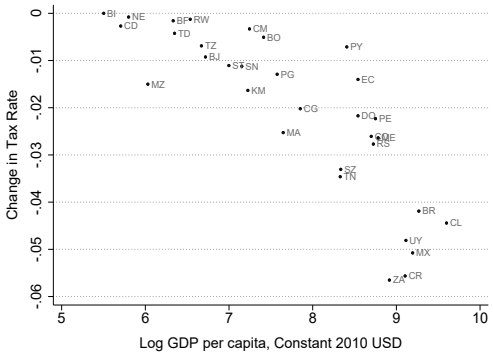
(b) Marginal Impact on Inequality



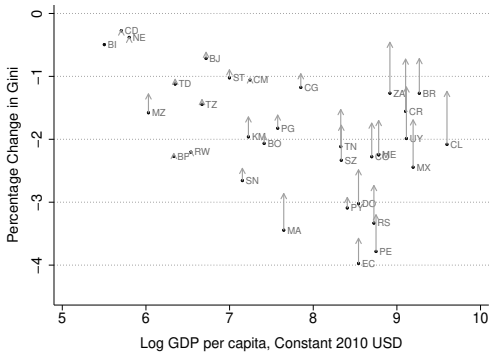
Notes: This figure shows the cross-country dispersion in optimal tax rates and inequality reduction achieved by setting optimal rates across the twelve main goods of the 2-digit UN COICOP classification. Panel (a) shows the dispersion in tax rates, measures as the normalized standard deviation of tax rates (coefficient of variation) for each country. Panel (b) shows the impact on inequality of further rate differentiation: starting from the Gini reduction achieved when differentiating across food and non-food goods, the arrows show the marginal increase of going to a full rate differentiation. More details in Section 7.2.

Figure A10: Impact of Income Taxes on Optimal Uniform Consumption Taxes

(a) Uniform Consumption Rate Change

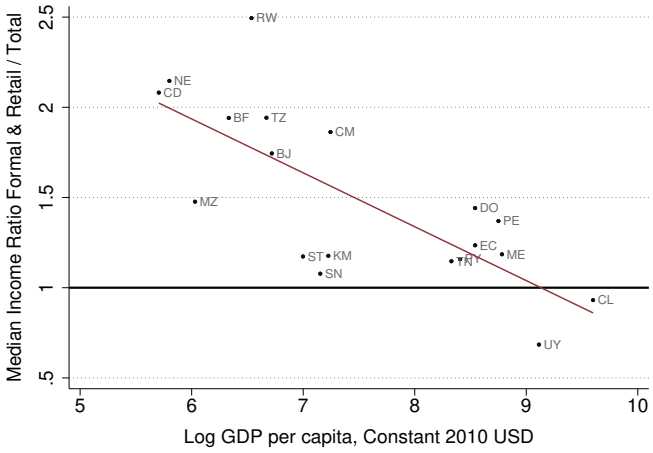


(b) Marginal Impact on Inequality



Notes: This figure shows how the presence of an income tax changes optimal uniform consumption tax rates, and the redistribution they can achieve, as a function of countries' log GDP per capita. The baseline is the uniform rates and inequality reduction achieved by consumption taxes without an income tax. In each panel, the arrow shows the marginal change upon incorporating an income tax. Panel (a) shows the change in uniform consumption tax rates. Panel (b) shows the marginal impact on inequality reduction from consumption taxes. More details in Section 7.2.

Figure A11: Income of Formal Retail Workers Relative to the Median



Notes: This figure plots the ratio of the median income of formal retail workers relative to all workers, as a function of countries' per capita GDP. Employment is formal if the worker contributes to health or pension funds. We conduct this analysis in the 19 countries of our sample where information on labor formality exists for the household head. More details in Section 8.2

.

Table A1: Main Reason for Choosing a Place of Purchase

	Outcome: Share of purchases (in %)		
Reason	**Informal Stores**	**Formal Stores**	**All Stores**
Access	42.1	31.3	41.5
Price	29.4	17.7	28.6
Quality	11.8	40.6	13.4
Store Attributes	6.9	5.0	6.9
Other	9.8	5.5	9.6

Notes: This table reports the share of purchases associated with different reasons, separately for purchases in informal and formal stores. Each number is an average across the six countries in our core sample in which the household survey contains these questions: Benin, Burundi, Comoros, Congo Rep., Morocco and RD Congo. In all surveys, seven reasons are listed which we classify into five categories as follows: access is defined as "The retailer is closer or more convenient" and "The good or service cannot be found elsewhere", price as "The good or services are cheaper", quality as "The goods or services are of better quality", store attributes as "The retailer offers credit" and "The retailer is welcoming or is a friend" and other as "Others reasons". More details in Section 4.2.

Table A2: Unit Values Across Places of Purchase

	Outcome: % dif. in formal sector unit values				# Purchases	# FE
Country	(1)	(2)	(3)	(4)	(5)	(6)
Benin	5.25	1.10	3.38	-0.39	262,280	5,065
	(7.10)	(5.66)	(7.53)	(6.19)		
Bolivia	4.08	3.53	4.69	3.86	120,971	1,549
	(1.40)	(1.12)	(1.40)	(1.15)		
Brazil	-0.11	-0.20	0.14	0.01	704,639	9,437
	(0.37)	(0.35)	(0.38)	(0.35)		
Burundi	2.53	4.39	4.81	5.23	250,139	2,454
	(4.65)	(4.73)	(4.39)	(4.23)		
Chad	-4.36	-3.21	-4.36	-3.21	380,462	1,968
	(1.80)	(1.77)	(1.80)	(1.77)		
Colombia	-0.33	-0.04	-0.30	-0.06	778,203	7,861
	(0.55)	(0.30)	(0.55)	(0.30)		
Comoros	22.56	14.93	21.81	14.49	113,228	1,142
	(5.01)	(3.64)	(4.98)	(3.64)		
CongoDRC	4.62	0.87	9.77	5.89	865,754	5,556
	(16.79)	(12.88)	(17.47)	(14.15)		
Congo Rep	27.84	23.70	27.12	23.01	208,557	1,182
	(5.88)	(4.67)	(6.03)	(4.78)		
Costa Rica	3.04	2.37	1.93	1.58	122,467	1,593
	(2.40)	(2.11)	(2.17)	(1.93)		
Dominican Rep	18.86	13.64	18.94	13.73	340,303	4,416
	(1.69)	(1.01)	(1.68)	(1.00)		
Ecuador	2.29	1.86	2.23	1.82	1,030,387	12,104
	(0.63)	(0.63)	(0.63)	(0.62)		
Eswatini	3.09	2.38	1.31	1.06	89,209	852
	(2.10)	(1.79)	(1.89)	(1.46)		
Mexico	1.10	1.00	1.10	1.00	446,417	6,195
	(1.16)	(1.02)	(1.16)	(1.02)		
Montenegro	10.36	9.57	7.13	6.45	138,446	867
	(3.70)	(3.25)	(3.08)	(2.85)		
Morocco	7.10	5.43	6.88	5.22	743,979	3,598
	(0.87)	(0.70)	(0.92)	(0.75)		
Peru	14.70	13.29	14.69	13.29	1,300,408	10,721
	(2.74)	(2.46)	(2.74)	(2.46)		
Sao Tome	6.81	4.87	6.69	4.86	215,527	2,946
	(1.39)	(1.37)	(1.39)	(1.34)		
Serbia	2.39	2.03	2.86	2.49	503,344	9,332
	(0.49)	(0.46)	(0.51)	(0.48)		
Tanzania	2.11	1.59	2.80	2.21	1,169,193	13,771
	(0.73)	(0.68)	(0.59)	(0.55)		
Avg. of 20 Countries	6.70	5.16	6.68	5.13		
Confidence Interval	[0.7,12.7]	[0.2,10.1]	[0.7,12.7]	[0.1,10.1]		
# of p-values < 0.05	12	12	11	11		
Winsorization [5,95]		X		X		
Self Consumption			X	X		

Notes: This table shows the percent difference in unit values between formal and informal stores. The sample is restricted to food purchases, where units and unit values are well measured, in the 20 sample countries with required data. We run the following specification: $ln(unit\ value)_{ipmu} = \beta\ Formal_{ipmu} + \mu_{pmu} + \epsilon_{ipmu}$, where $ln(unit\ value)_{ipmu}$ is the unit value reported by household i, for product p, in location m, in units u, and $Formal_{ipmu}$ equals one if the product is purchased in a formal store. We add fixed effects at the level of [product]*[location]*[unit]. Standard errors are clustered at the location level. More details in Section 4.2.

Table A3: Ratio Top over Bottom Quintile of Effective Tax Rates

Tax policy	Baseline Assignment	Country-Specific Assignment	Baseline + Distributional Savings	Baseline + Mexican Reform
Uniform rate, only formal taxed	2.03	1.77	1.94	1.63
Food exempt, formal and informal taxed	2.27	2.14	2.17	2.08
Food exempt, only formal taxed	1.64	1.64	1.56	1.64

Notes: This table shows the progressivity of consumption tax policies, measured as the ratio of the effective tax rate paid by the richest household quintile over that of the poorest quintile. The numbers are averages for the 32 sample countries. The rows correspond to the three tax policy scenarios: (1) a uniform tax rate on all goods when only formal goods are taxed, (2) a zero rate on food, when only formal goods are taxed, (3) a zero rate on food, in the unrealistic setting where both formal and informal goods are taxed. Each column corresponds to a different set of assumptions. Column (1) corresponds to the baseline informality assignment (modern retailers are formal and traditional retailers are informal). Column (2) assigns country-specific probabilities of formality to modern and traditional stores, based on the country retail reports and data from the World Bank Enterprise Surveys. Column (3) applies the tax pass-through for informal stores (14%) and formal stores (77%) estimated from the VAT reform in Mexico, to all countries in the sample. Column (4) adds savings rates which are decile and country specific, using data from Global Findex. More details in Section 5.3.

Table A4: Change in Gini from Optimal Tax Policy

Tax policy	(1) Baseline Assignment	(2) Cross-variety Elasticity	(3)	(4) Distributional Savings	(5) Country-Specific Assignment	(6) Mexican Reform Pass-through
Panel A: Average across countries		$\tilde{\epsilon}^C = 1$	$\tilde{\epsilon}^C = 2$			
Uniform rate	-1.92	-2.34	-1.63	-1.84	-1.73	-1.23
Food rate differentiation	-2.64	-3.10	-2.31	-2.53	-2.39	-1.85
Full rate differentiation (12 goods)	-3.20	-3.56	-2.84	-3.00	-2.86	-2.32
Uniform rate with PIT	-1.56	-1.91	-1.32	-1.43	-1.39	-1.01
Food rate differentiation with PIT	-2.12	-2.51	-1.86	-2.04	-1.92	-1.51
Panel B: Middle/Low Income Countries Ratio						
Uniform rate	1.52	1.36	1.65	1.50	1.51	1.37
Food rate differentiation	1.79	1.57	1.98	1.75	1.81	1.61
Full rate differentiation (12 goods)	1.93	1.70	2.11	1.87	1.94	1.72
Uniform rate with PIT	1.26	1.13	1.36	1.29	1.24	1.14
Food rate differentiation with PIT	1.50	1.31	1.66	1.52	1.51	1.35

Notes: This table shows the redistributive impact of different consumption tax policies under different hypothesis. Our metric for redistribution is the percent change in Gini from the pre-tax income distribution to the net-of-tax distribution. We take the average across the 32 countries in the sample in the first panel of the table, and compare middle income to low income countries in the second panel, by taking the ration of their respective Gini reductions; for example a ratio of 2 implies that this policy reduces the GINI twice as much in middle income compared to low income countries. The rows correspond to the tax policy scenarios considered: (1) uniform rate on all goods (2) optimal differentiated tax rates on food and non food goods (3) optimal differentiated tax rates for each of the 12 large goods categories (COICOP 2 digit) (4) uniform rate on all goods with an actual personal income tax (5) optimal differentiated tax rates on food and non food goods with an actual personal income tax. The columns correspond to the different assumptions about the data. The baseline in column (1) corresponds to the central assignment of modern retailers to VAT-formality, of informal retailers to full informality, and a value of the cross-variety elasticity of substitution between formal and informal varieties of 1.5. Columns (2) and (3) vary the elasticity of substitution between its plausible bounds from 1 to 2. Column (4) adds country-specific distributional savings rates. Column (5) assigns the country-specific probabilities of formality status to modern and traditional stores, based on the Euromonitor country reports and the World Bank Enterprise Surveys. Column (6) assumes a 14% pass-through of taxes to traditional stores and 77% pass-through to modern stores, based on the estimates from the Mexican VAT reform.

B Appendix: Data on Consumption by Store-Type

All codes to replicate the paper are available at https://github.com/pierrebachas/Informality_Taxes_Redistribution. This includes cleaning files for each country's micro data, and all files generating tables and figures in the paper.

Inclusion Criteria Our dataset consists of 32 nationally representative household budget surveys. We use surveys which satisfy the following three criteria:

1. The expenditure module(s) in the survey is structured as an open consumption diary, rather than a pre-filled diary covering a limited set of products.
2. The expenditure survey includes a variable for the place of purchase (data on where each item was purchased). The place of purchases are detailed enough for us to apply our taxonomy of store types, as outlined below.
3. The place of purchase variable rarely contains missing values, particularly for food, clothing and household goods product categories (see Figure A1).

Data Sources and Coverage Most of our data come from two main sources: (i) the World Bank Microdata Library and (ii) National Statistical Agencies. To access the data we searched the restricted access World Bank Microdata Library for household expenditure surveys for which the above criteria appeared to be satisfied. The surveys which satisfied these criteria varied in their ease of access: for some countries, the micro data were accessible for download on the World Bank platform, others were licensed and required applications through the World Bank, which would in turn sometimes contact the country's national statistical agency for approval. If a survey was listed without its micro data through the World Bank platform, we reached directly the country's' statistical agency.

The countries which satisfied the criteria for inclusion span four regions of the world, concentrated in Sub-Saharan Africa and Latin America and the Caribbean, as detailed in Table B1. Unfortunately we could not include countries in Asia—except for Papua New-Guinea—since the question on the place of purchase was always missing from their budget surveys. The online appendix lists the 32 countries which we include, with summary statistics on the structure of each survey. Any slight deviation from our inclusion criteria is outlined.

Table B1: Regional Survey Representation

Region	# Countries	Pop. of Surveyed Countries (Millions)	Total Pop. (Millions)	Proportion of pop.
Sub-Saharan Africa	16	379	1078	35%
Middle East & North Africa	2	48	449	11%
Europe & Central Asia	2	9	918	1%
Latin America & Carribean	11	496	641	77%
East Asia & Pacific	1	9	2328	0.4%

Consumption Module Structure Surveys are not homogeneous across countries. We provide a summary below (country by country information in the online appendix):

- **Number and frequency of modules:** the number of consumption modules ranges from 1 to 17 across countries in the sample. All surveys have a module which is a diary of consumption over some short to medium period of time and some countries complement these with recall modules for more infrequent purchases. For example, Costa Rica has a single consumption module, while Morocco has 17, with modules specialized by frequency and products. Surveys with multiple modules typically asked for consumption linked to the frequency of expenditures (e.g. monthly, quarterly).

- **Durables:** durable items are included whenever available, but their inclusion is more probable in surveys which have recall modules.

- **Home-production:** home production is included as a "place of purchase" for all countries but Chile where it is not available. In some countries, it was precoded as an option for the place of purchase, while in other cases we added it as a place of purchase based on other variables, such as "acquisition mode" which had "purchased or "self/home production." Self-production purchases are typically valued using the local market value.

- **Product codes:** all surveys have product codes for each consumption item, which typically follow the United Nations Classification (COICOP) or which we can match to the COICOP with a cross-walk. For a few countries we could not find a product crosswalk. We used each country's own classification scheme' for these four countries (Brazil, Chad, Peru and Tunisia).

Categories of Places of Purchase Each of the 32 countries have a place of purchase for each transaction. The different place of purchases options available to respondents however differ across countries. However, these can be classified into broad categories which are roughly equivalent across countries. To the best of our knowledge, the only other project which constructs a common taxonomy of places of purchases across countries is the International Price Comparison (ICP) project, which builds purchasing power parity indexes. The ICP provides a store type classifier for marketed consumption which is used by individual countries to obtain price quotes from a variety of retailer types. Our classification mirrors that of the ICP. Consumption of goods is categorized into five broad categories of places of purchase: (1) non-market consumption (e.g. home-production); (2) Market consumption, no store front (e.g. markets, street stalls); (3) Market consumption, corner and convenience shops; (4) Market consumption, specialized shops (e.g. brand stores); (5) Market consumption, large stores (e.g. supermarkets, malls). Consumption of services is categorized into four categories: (6) Services provided by institutions (e.g. bank, hospital, university); (7) Service provided by individuals (e.g. maid services, gardening); (8) Entertainment (e.g. restaurants, hotels); (9) Informal Entertainment (e.g. food truck).

The majority of countries have places of purchase for each of the five good categories. Four countries do not distinguish between specialized stores (category 4) and corner/convenience stores (category 3). For these countries, we use the following methodology: i) for each decile we compute the budget share of categories 3 and 4 together using the survey. ii) for each decile we compute the average share of category 3 in the total budget share of categories 3 and 4 in comparable countries, where we define 'comparable' as countries whose average GDP per capita is in the 50-150% range of the country's GDP per capita. iii) We then impute the relative shares of categories 3 and 4 in the country using these average relative shares at the decile level.

In some countries all services do not have a specified place of purchase. In particular there is no detailed list of institutions as potential places of purchases. These are also countries in which the share of expenditures with an 'unspecified' place of purchase is larger: indeed when looking at what types of products compose the unspecified category, over half are utilities, while the remaining is mainly education and health spending. Finally, we assign the remaining places

of purchase that are harder to categorize (e.g. purchases from internet or from abroad) to category (6) "services provided by institutions". We note that these represent less than 0.5% of total expenditure.

For replication purpose the countries-specific assignment of specific places of purchase to the broad categories presented above is detailed in the online appendix, where each place of purchase representing more than 0.5% of total expenditure is included with its original name in the survey.

C Appendix: Formality Assignment of Store Types

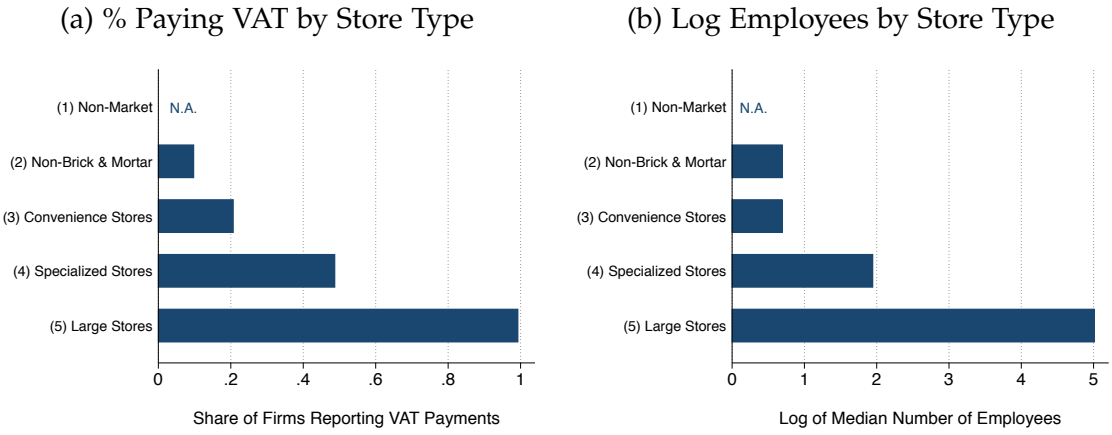
C.1 Evidence on Sales Formality Status by Store-Type Across Countries

Mexico's Firm Census The 2013 Mexican firm census details store-type classification for the universe of modern and traditional retailers in the country – with the caveat that home-based vendors and makeshift stalls that change location frequently are excluded. The absence of data on these store-types likely leads us to overestimate the sales formality status among all traditional retailers. The census also asks firms for their remittance of VAT on their sales and on their inputs. We define a store to be VAT formal if it levies VAT on any of its sold products. This data allows us to directly observe VAT status by detailed store-type. Figure C1 Panel (a) shows that the share of traditional retailers (non-brick and mortar; convenience stores) remitting VAT is 10%. This contrasts with modern stores, where the VAT-share is between 51% (in specialized stores) and 98% (in large stores such as supermarkets). Traditional and modern retail stores therefore differ substantially in the extent of VAT sales formality. This is likely driven in part by the large size difference between these two types of stores: C1 Panel (b) shows a large difference in the median number of employees. This increases the likelihood that traditional retailers are legally exempt from VAT and constrains enforceability.

Cross-country Evidence: Data and Methodology How does the sales formality status of traditional retailers in Mexico compare to other countries at different income levels? We combine two datasets used in the literature, described below.

The first source is the World Bank firm surveys: the Enterprise Surveys (ES) and the Informal Surveys (IS). 35 low and middle income countries have both of these survey types (from Africa, Asia and Latin America). These surveys contain

Figure C1: Formality by Store Type in Mexico



For construction of these graphs, see Section C.1.

harmonized information for firms in the retail sector on registration status with central government agencies, which we use as the measure of formality.[37] The Enterprise Survey samples firms from geographical areas which are likely to contain a high number of formal firms with more than 5 employees. The Informal Survey samples from areas which are likely to have a high concentration of small and informal firms. Once these sampling zones are established, all firms located in the zone, both registered and unregistered, are considered for the final sample. We refer to La Porta and Shleifer (2014) for a detailed discussion of the sampling methodology and formality definition. Combined, the ES-IS sampling frames provide data-coverage of firms in the traditional and modern retail segment.[38]

The second source are the retail market reports produced by Euromonitor International, which are described in Section 2. These reports contain detailed information on sales and number of units by harmonized store type categories of modern and traditional grocery retail, which we will consider as the retail sample. Bronnenberg and Ellickson (2015) use these in their review of global retail patterns. The modern and traditional store types are consistent with our classification, with the exception that home production is not included in the

[37] For most surveys, registration refers to the central tax administration (in charge of the VAT). We include the remaining surveys to retain as broad a sample as possible; other agencies that firms can be registered with include the chamber of commerce and the national statistics office.

[38] The main exception is that IS cover only incompletely the smallest retailers, including home producers and 'street hawkers', leading us to overstate the formal share in all traditional stores.

Euromonitor reports, while we assign it to traditional stores.

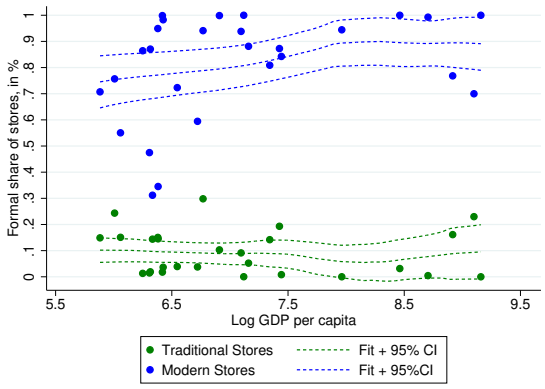
Based on the WB survey data, we can measure the formality probability for retail firms as a function of sales, but not by store type. Based on the Euromonitor data, we can measure average sales by store type (modern or traditional), but not by formality status. We combine these two data sources to obtain a plausible estimate of the average formal share of traditional and modern retailers. To increase the usefulness of combining sales information from different data, we use Euromonitor sales from the same year as the surveys. We also create weights which ensure that the WB based sales distribution is representative of the universe of retailers captured in the Euromonitor data. Despite this, concerns remain about the extent of direct comparability of the retail sales values between ES-IS and Euromonitor. Therefore, our preferred approach is to calculate the ratios of traditional and modern store sales to average retail sales from Euromonitor, and use these ratios to impute the average traditional and modern store values in the ES-IS sample. We define the traditional (modern) formal share as the formality share at which the formality distribution by size intersects the average traditional (modern) size. Ideally, we would superpose a density size distribution and calculate the store type formality share by integration. Unfortunately, the Euromonitor data only allows us to measure averages by store-type.

Cross-Country Results The modern and traditional formality shares are plotted against log GDP per capita for the 33 ES-IS countries in Figure C2, panel (a). Two results emerge: first, the average formal share is 5-15 percent in traditional stores and 75-90 percent in modern stores. These shares are comparable to those obtained in Mexico using retail Census data. Second, the average formality share of traditional stores is stable across countries, and that of modern stores only slightly increases with development.

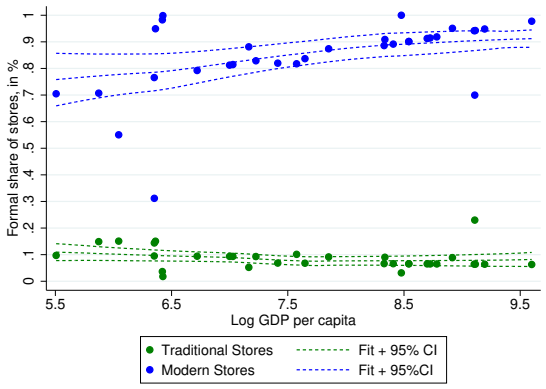
We use the constructed variables to create a data driven, country-specific formality assignment to modern and traditional stores (panel (b) Figure C2). In the 11 countries where we have ES-IS data and expenditure surveys, we directly use the measured values. For the remaining 21 countries in our database, we use the predicted formality shares based on a regression of the modern (traditional) formality share against log GDP per capita and continent dummies. Even though the sample size is small, the relative stability of formality shares across log GDP

Figure C2: Formality by Store Type Across Development

(a) Formality Share by Store Type



(b) Formality Share in our Sample



For construction of these graphs, see Section C.1.

per capita and continent suggests that this exercise yields meaningful predictions.

Relation to VAT exemption thresholds We record the sales value of the VAT exemption threshold for all countries in our data. Data on VAT thresholds comes from the harmonized set of country tax code reports produced by the International Bureau of Fiscal Documentation. We follow the methodology described in Keen and Mintz (2004). In most countries outside Latin America, the value of the exemption threshold is explicitly defined. In many Latin American countries, the threshold value is the level of sales below which firms are registered for a simplified tax regime rather than VAT. The average size of traditional and modern retail stores, along with value of the VAT exemption threshold, are reported in the online appendix. We find that the ratio of sales-values to the VAT exemption threshold is 1.01 for traditional stores and 38.85 for modern stores.

C.2 Evidence from Mexican VAT Reform

We exploit a 2014 reform in Mexico to estimate the pass-through of a VAT rate increase to consumer prices separately for modern and traditional retailers. We rely on the monthly microdata collected by the national statistics office, INEGI, to construct the measure of consumer price index (CPI) in Mexico. INEGI enumerators obtain price quotes for approximately 83,500 items that cover 315 product

categories in 141 municipalities. The price sampling is meant to be representative of Mexican household consumption and contains, importantly for our purposes, price quotes from both modern and traditional retail stores (including street stalls and other non-brick and mortar stores).

We obtain access to the confidential data-set of the Mexican CPI, which allows us to observe the municipality location as well as uniquely identifying information on the store at which the price quote was collected. We can merge this data with the 2013 Mexican census at the unique store level, in order to obtain detailed information on the actual tax status of the store. This allows us to construct measures of product level prices on a monthly basis in both informal and formal retail stores (based on VAT remittance status) and verify that our results are robust to estimating pass-through by actual tax-status (rather than by store-type). This robustness is due to the strong overlap between tax-status and store-type (Figure C1). We focus on the 127 products which constitute the core inflation items and which are subject to VAT. In addition, we focus on non-tradeable products due to concerns over bias that would arise from indirect treatment in control (non-border) areas. Estimation is at the product-location-month price level.

We leverage the variation in VAT rate induced under a reform passed by the Mexican government in January 2014. Prior to the reform, border areas benefited from a VAT reduction, such that the VAT rate was 11% rather than the standard, non-border 16%. The reform increased the VAT rate in the border areas with immediate effect from 11% to 16%, while keeping rates unchanged in the non-border areas. The reform was motivated by the government's objective to equalize tax incentives across the country.

We use a difference-in-differences (DiD) design to estimate the VAT rate increase pass-through to consumer prices in informal stores. The DiD coefficients are in Figure 9. The main results are based on a sample without restrictions, but are robust to excluding observations from Mexico City; restricting control non-border areas to be geographically close to border areas with price data; and, winsorizing 1% of price observations (results available upon request).

D Appendix: Theory And Calibration

D.1 Proofs of model statement

Efficiency cost of taxation in the presence of an informal sector The efficiency cost of taxing $j1$, the formal variety of product j, is determined by the uncompensated elasticity of demand for the good, ϵ_{j1}.

We start by writing the uncompensated elasticity of product j as a function of the uncompensated elasticities of varieties $j1$ and $j0$ and the cross-variety price elasticities, considering a price change such that $\partial p/p = \partial p_1/p_1 = \partial p_0/p_0$:

$$\epsilon_j = \epsilon_{j1}(1-\alpha_j) + \epsilon_{j0}\alpha_j + \epsilon_{j1,0}(1-\alpha_j) + \epsilon_{j0,1}\alpha_j \tag{7}$$

where $\alpha_j = \frac{p_0 x_0}{px}$ is the share of informal consumption in total consumption of the product and $\epsilon_{j0,1}$ ($\epsilon_{j1,0}$) it the elasticity of demand for the informal (formal) variety with respect to the price of the formal (informal) variety.

Writing ϵ_j^C the compensated price elasticity of product j, the Slutsky equation gives $\epsilon_j = \epsilon_j^C - \eta_j s_j$. Decomposing income elasticities η as $\eta_j = \eta_{j1}(1-\alpha_j) + \eta_{j0}\alpha_j$, using $s_j = s_{j1} + s_{j0}$ and replacing we obtain:

$$\epsilon_j^C = \epsilon_{j1}^C(1-\alpha_j) + \epsilon_{j0}^C\alpha_j + \epsilon_{j1,0}^C(1-\alpha_j) + \epsilon_{j0,1}^C\alpha_j \tag{8}$$

Slutsky symmetry implies $\epsilon_{j1,0}^C x_1/p_0 = \epsilon_{j0,1}^C x_0/p_1$. Writing $\epsilon_{j0,1}^C = \tilde{\epsilon}_j^C$, using our assumption of equal compensated own-price elasticity across varieties within products ($\epsilon_{j1}^C = \epsilon_{j0}^C, \forall j$) and re-arranging, we obtain:

$$\epsilon_{j1}^C = \epsilon_j^C - 2\tilde{\epsilon}_j^C\alpha_j \tag{9}$$

To obtain an expression for the compensated price elasticity ϵ_{j1}, the parameter of interest, we use the Slutsky equation again and obtain:

$$\epsilon_{j1} = \epsilon_j^C - 2\tilde{\epsilon}_j^C\alpha_j - \eta_{j1}s_{j1} \tag{10}$$

The uncompensated elasticity of demand for the formal variety is therefore increasing (in absolute value) in the elasticity of substitution across varieties, $\tilde{\epsilon}_j^C$, and the share of the informal variety in total consumption of the product.

The parameter ϵ_{j1} defined in (10) is the efficiency parameter in the expression

for the optimal rate on product j, equation (5) in the text. To go from the product level elasticity ϵ_{j1} to the elasticity for all formal varieties, ϵ_1 in (4), we use:

$$\epsilon_1 = \sum_j \epsilon_{j1} \frac{x_{j1} p_{j1}}{x_1 p_1} \tag{11}$$

where p_1 is the aggregate price for all formal varieties.

Optimal rate on product j with non-null cross-goods elasticities of substitution In this section we relax the assumption that $\epsilon_{j1,k1}$, the elasticities of substitution across formal varieties of goods, are always equal to zero. Taking the derivative of (3) with respect to t_j and re-arranging we obtain a new version of expression (5), the optimal rate on good j:

$$\tau_j^* = \frac{\int_i (\bar{g} - g^i) \phi^i \frac{s_{j1}^i}{s_{j1}} di}{-\bar{g}(\epsilon_{j1} + \sum_{k \neq j} \epsilon_{k1,j1} \frac{x_{k1} q_{k1} t_k^*}{x_{j1} q_{j1} t_j^*})} \tag{12}$$

Where $\epsilon_{k1,j1} = \frac{\partial x_{k1}}{\partial p_{j1}} \frac{p_{j1}}{x_{k1}}$ is the elasticity of demand for formal variety $k1$ with respect to the price of the formal variety $j1$. The additional term in the denominator reflects how the cross-goods price effects affect the efficiency cost of taxing the formal variety of good j. When all other formal varieties are (uncompensated) substitutes ($\epsilon_{k1,j1} > 0, \forall k1$) the efficiency cost is lower, because increasing the rate on j increases consumption of other formal varieties and therefore the taxes collected on these varieties. When all other formal varieties are (uncompensated) complements ($\epsilon_{k1,j1} > 0, \forall k1$) the efficiency cost if higher, because increasing the rate on j decreases consumption of other formal varieties and taxes collected on those. When we calibrate the rates τ_j^* we only consider large good categories: food and non-food goods for our baseline results, and 12 large product categories (food, textiles, health, etc) as an extension. Because these categories are large the cross-price effects are likely to be very small relative to the own-price effects.

D.2 Calibration

Calibration details Here we explain how we calibrate the optimal tax rates defined in expressions (4) and (5). Table 3 summarizes our choice of calibration

parameters. First, we calibrate several parameters directly from our data: we use the observed budget shares shown in Section 4, household expenditure to proxy for household income and the the informal shares of consumption for each good and country. We relax our assumptions that Engel curves are log-linear and that development does not affect within-country inequality, using instead the observed budget shares and income distributions in each country.[39]

We use our data to obtain estimates of income elasticities for all goods and varieties. To obtain the income elasticity of demand for the formal variety, η_{j1} we use our estimates of the slope of the formal Engel curve for good j, β_{j1}, and the expression $\eta_{j1} = 1 + \frac{\beta_{j1}}{s_{j1}}$. We obtain income elasticities η_j using $\eta_j = 1 + \frac{\beta_j}{s_j}$.

Second, we use existing literature to calibrate the remaining parameters. There are no estimates of the cross-price compensated elasticity of demand between formal and informal varieties $\tilde{\epsilon}^C \alpha_j$ so we use estimates of the elasticity of substitution in consumption across store-types. The cross-price elasticity is related to this elasticity of substitution σ in a CES utility function by the expression $\epsilon^C_{0,1} = \sigma s_0$ where s_0 is the share of informal consumption of total consumption of the good. Faber and Fally (2017) estimate an elasticity of substitution between large and small stores in the US of 2.2, Atkin et al. (2018b) estimate the elasticity of substitution between foreign and domestic supermarkets and find estimates in the 2-4 range. We therefore use 3 as our baseline of σ. For an average value of s_0 of 0.5 this yields a baseline value of $\tilde{\epsilon}^C \alpha_j$ of 1.5, we consider the range 1-2 as a robustness check. We set a value of -0.7 for the own-price compensated elasticity of goods. Together, these parameters yield values for the own-price uncompensated elasticity of goods (calibrated using expression (14) that are in the $[-2, -0.5]$ range, in line with estimates from the literature (see for example Deaton et al., 1994).

Finally, we specify government preferences by setting the same social welfare weight for households in a given decile of the household expenditure distribution in each country. The welfare weights are obtained to match an average uniform tax rate of 18%. This implies that governments place eleven times more weight on income received by households in the poorest decile than in the richest decile.

[39] Note that our model calls for using budget shares observed under a counterfactual 'no tax or transfers' scenario. We do not attempt to adjust observed budget shares to take into account the fact that they are affected by current tax systems as this is beyond the scope of this paper.

Role of inequality: optimal uniform rates with no informal sector Income inequality affects optimal rates through the $\phi^i = \frac{y^i}{\bar{y}}$ term. With standard welfarist preferences higher inequality increases the equity gain from levying progressive taxes, and therefore tends to increase tax rates. To illustrate the (small) role played by inequality in our results Figure A5 plots optimal uniform rates obtained under the assumption that all varieties can be taxed with the same uniform rate - there is no informal sector. The only difference across countries in this policy scenario is the level of inequality, the distribution of the ϕ^i term. Recall that we use household expenditures per capita to proxy for household income, and that we are unlikely to observe incomes at the very top of the income distribution because we use household expenditure surveys in which the very richest households are known to be under-represented.

Figure A5 shows that inequality does vary across countries - the optimal uniform rates in the absence of an informal sector vary from 28% to 55% - but there is no systematic relationship between total expenditures inequality and economic development: the best linear fit line is extremely flat. The relationship between economic development and optimal linear rates when only formal varieties can be taxed, seen in Figure 7 is therefore not due to a systematic change in inequality over development but is explained by the changes in patterns of formal consumption across countries.

Changing government preferences Figure A8 presents results obtained when varying government preferences. Our baseline results assume that the government values one extra dollar in tax revenues as much as one extra dollar of income distributed equally across households, so that $\mu = \bar{g}$, where $\bar{g}$ is the average marginal social welfare weight. The top panel of Figure A8 assumes instead that the government values tax revenues more than the average marginal social welfare weight, so that $\mu = 1.1\bar{g}$, the bottom panel assumes that it values tax revenues less than the average marginal social welfare weight, $\mu = 0.9\bar{g}$.

We see that, as expected, a higher value for μ increases tax rates: the average optimal uniform rate is 16% for the low value of μ, 20% for the high value, and 18% for our baseline. Increasing μ increases the optimal rate on food slightly more than that on non-food: the optimal relative food rate (ratio of optimal food rate to optimal non-food rate) is 0.71 for the high value of μ, 0.67 for the low

value and 0.69 for our baseline. Patterns for change over economic development are unaffected.

D.3 Personal Income Tax Extension

Model extension This section models optimal consumption tax policy in a context in which the government also levies a personal income tax (PIT). The PIT parameters, assumed exogenous, are an income threshold $\bar{y}$ above which households pay an income tax on their income $y^i - \bar{y}$, and a tax rate t_y. Household i's disposable income y^i_D after paying their PIT (but prior to any consumption taxes) can therefore be written as $y^i_D = (y^i - \tilde{y})(1 - t_y) + \tilde{y}$ if $y^i > \bar{y}$ and $y^i_D = y^i$ otherwise. The government sets optimal tax rates t_j on each good j to maximize:

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

The PIT affects the indirect utility function, which is now a function of disposable post PIT-income y^i_D and household expenditure on each good x_{j1}, which is also (implicitly) a function of y^i_D. The first-order-condition for $\tau_j = \frac{t_j}{1+t_j}$ is:

$$\tau_j = \frac{\int (\mu - g_i) p_{j1} x^i_{j1}}{-\mu \epsilon_{j1} p_{j1} x_{j1}}$$

where both g_i, household i's marginal social welfare weight, and x^i_{j1} are now defined as a function of disposable income y^i_D. The existence of a PIT increases the welfare weight of households above the PIT threshold $\bar{y}$; all else equal this tends to decrease the optimal rate on goods consumed more by richer households, for example formal varieties of all goods. It decreases the consumption of the formal variety of good, $j1$, the more so the higher the income effect η_{j1}; this also tends to lower the optimal rate on goods consumed more by richer households. Over development Jensen (2019) documents a fall in the PIT threshold $\bar{y}$. This lowers the disposable income of richer households, by increasing taxes on existing taxpayers (increasing $y^i - \bar{y}$) and pushing new households into the tax net.

Calibration with Incomes Taxes We obtain the PIT parameter $\bar{y}$ for each country from Jensen (2019) where available. For the 10 countries for which no data is

available, we interpolate the parameters using the observed relationship between these parameters and each country's log GDP. The parameter $\bar{y}$ is expressed as a percentile of the household income distribution. We assign $t_y = 50\%$ for all countries in our sample. This corresponds to the largest observed top marginal tax rate across the sample of countries in Jensen (2019). This choice is conservative, as it overstates the redistribution achieved by the PIT, and thus understates the redistribution potential of consumption taxes. Jensen (2019) also documents that the top marginal rate does not significantly vary across countries with different income per capita.

We then compute each household's disposable income y_D^i using the definition given above and total expenditure per capita to proxy for income y^i. From this, we obtain each households' consumption x^i on each good and variety as a function of its disposable income: for example $x_{j1}^i(y_D^i) = x_{j1}^i(y^i) + \eta_{j1}\frac{x_{j1}^i(y^i)}{y^i}(y_D^i - y^i)$, where $x_{j1}^i(y^i)$ is household i's expenditure on the formal variety of good j observed in the data. We use the estimates of income elasticities obtained using our data, see table 3.

To calibrate the change in the marginal social welfare weight g^i due to a PIT, we specify $g^i(y^i) = \frac{-\alpha}{y_i}^2$. This functional form is what we would obtain if we specified a log indirect utility function $v()$ and a utilitarian social welfare function $G()$. We calibrate it to obtain the distribution of the g^i terms of our baseline specification (Table 3). The new weight g^i is then given by $g^i(y_D^i) = g^i - \frac{\alpha}{y^{i2}}(y_D^i - y^i)$.

D.4 VAT pass-through within supply chains

In this subsection, we show that our assumptions in the main formality assignment regarding the pass-through of taxes to prices in the formal and informal sector can be modelled as equilibrium responses of firms in a simple supply side model. Each variety $j1$ is produced by a firm that pays taxes (formal firm), and each variety $j0$ by a firm that does not pay taxes (informal firm). This assigns firms to a formality status based on whether taxes are levied on sales. We assume that the tax takes the form of a VAT.

Upstream firms k use only labor $x_k = L_k$. Downstream firms produce varieties jl using inputs produced by upstream firms and have the production function:

$$x_{jl} = \Big(\sum_k \alpha_{jlk} x_{jlk}^{\frac{\rho-1}{\rho}} \Big)^{\frac{\rho}{\rho-1}} \tag{13}$$

where x_{jlk} is the quantity of inputs k used by the downstream firm producing variety jl, and ρ the constant elasticity of substitution in production. Labor is paid a fixed wage w. The value of sales for the downstream firms is given by $q_{jl}x_{jl}$ where q_{jl} are the endogenous producer prices, which then determine consumer prices: $p_{j1} = q_{j1}(1 + t_j)$ if the firm is formal, $p_{j0} = q_{j0}$ if the firm is informal. We assume firms compete under monopolistic competition, which implies they maximize profit π_{jl} whilst accounting for the demand function $x_{jl}(p_{jl})$ they face. Writing ϵ_{jl} the price elasticity of demand for variety jl and taking first-order condition with respect to prices, we obtain the expression for consumer prices:

$$p_{jl} = (1 + t_j f_{jl}) \frac{P_{jl}}{\phi_{jl}} \frac{\epsilon_{jl}}{\epsilon_{jl} - 1} \tag{14}$$

where $f_{jl} = 1$ if the firm producing jl is formal, zero otherwise, and P_{jl} is its input cost index. P_{jl} is obtained by cost minimization and equal to:

$$P_{jl} = \Big(\sum_k \alpha_{jlk}^{\rho} p_{jlk}^{1-\rho} \Big)^{1/(1-\rho)} \tag{15}$$

Here p_{jlk} is equal to the net of tax price paid for the product k by the firm producing variety jl. We assume, under a VAT, that: no tax is paid if both firms k and jl are informal; no tax is paid if firm k is informal; tax is paid on the transaction only if upstream firm k is formal and firm jl is informal. Formally:

$$p_{jlk} = (1 + t_k f_k (1 - f_{jl})) w \frac{\rho}{\rho - 1} \tag{16}$$

Combining expressions (14), (15) and (16), we can write the pass-through of taxes to the price of formal and informal downstream firms. The pass-through of taxes to the price of formal downstream firms ($f_{jl} = 1$) is equal to 1:

$$\frac{\partial p_{j1}}{\partial t_j} \frac{1 + t_j}{p_{j1}} = 1 \tag{17}$$

The pass-through of taxes to the price of informal downstream firms ($f_{jl} = 0$) can be written as:

$$\frac{\partial p_{j0}}{\partial t_j} \frac{1 + t_j}{p_{j0}} = s_{j0F} \tag{18}$$

where s_{j0F} is the share of formal inputs in firm $j0$'s total production costs:

$$s_{j0F} = \sum_k f_k \alpha_{j0k}^{\rho} P_{j0}^{\rho-1} p_{j0k}^{1-\rho} \tag{19}$$

The model shows that the pass-through of a tax rate increase via formal suppliers to informal retailers will be equal to the formal input share in total costs.

Inputs of Informal Retailers An ideal database to measure this input share would consist of retail Censuses across countries which contain detailed information on input sourcing by store-type. Amongst the countries in our sample, we were only able to locate the required information in the Mexico Retail Census. In this data-set, among informal retailers, only 8% report paying VAT on inputs, which applies on average to 40% of their intermediate purchases. The informal retailers that report positive VAT on inputs account for 25% of all informal sales. Combined, these data yield an overall formal input cost-share of 10% for informal retailers in the Mexican Census. Note that home producers are not included in the retail Census. If home producers are plausibly less likely to source inputs from formal suppliers than the other informal retail categories, the Census based estimate of formal input-share will be an upper bound for the input share of all informal retailers.

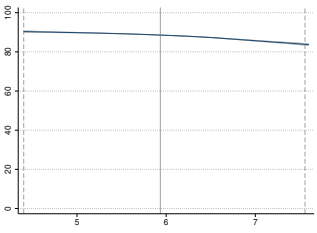
Supplementary Online Appendix, Not for Publication "Informality, Consumption Taxes and Redistribution"

by Pierre Bachas, Lucie Gadenne and Anders Jensen

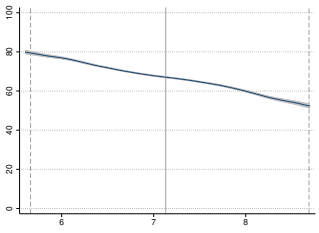
List of Figures

List of Tables

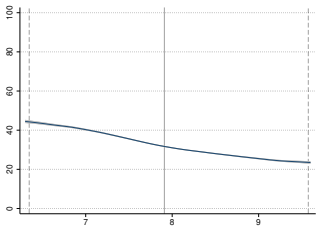
Figure E1: Informality Engel Curves



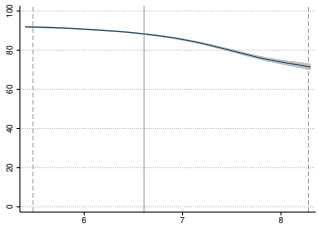
(a) Benin



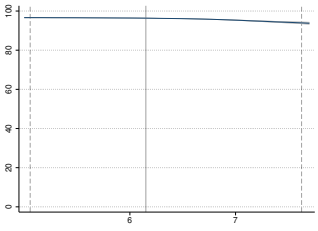
(b) Bolivia



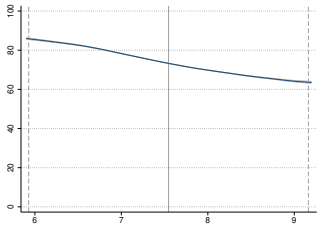
(c) Brazil



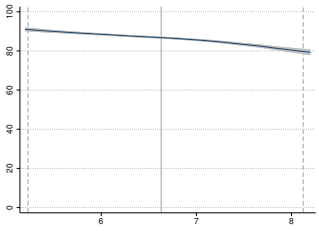
(d) Burkina Faso



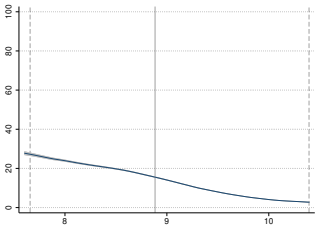
(e) Burundi



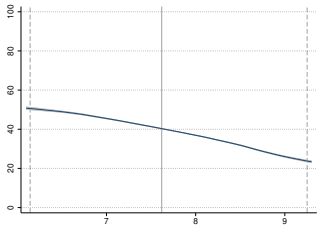
(f) Cameroon



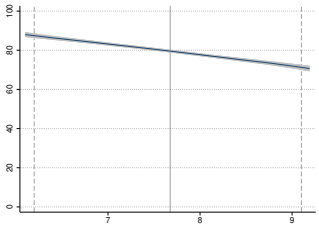
(g) Chad



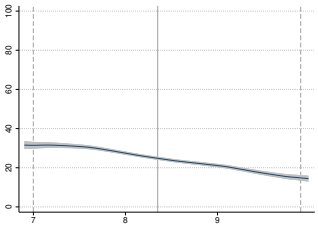
(h) Chile



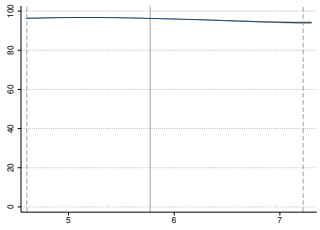
(i) Colombia



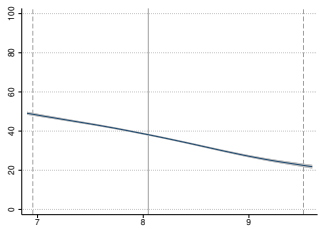
(j) Comoros



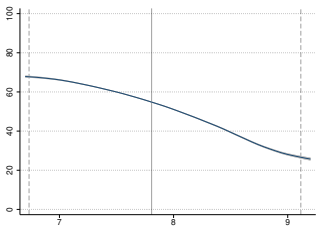
(k) Costa Rica



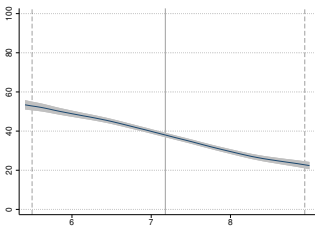
(l) Dem. Rep. of Congo



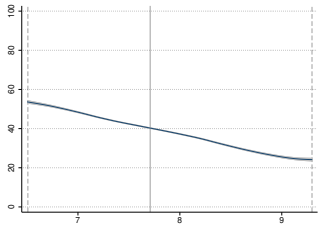
(m) Dominican Rep.



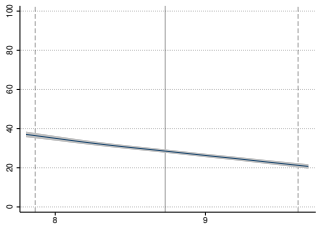
(n) Ecuador



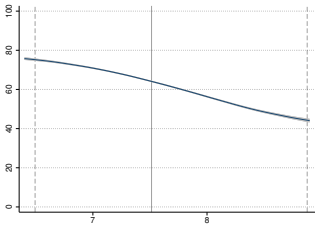
(o) Eswatini



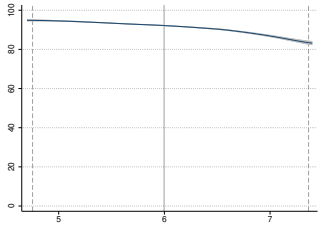
(p) Mexico



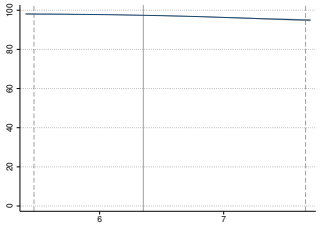
(q) Montenegro



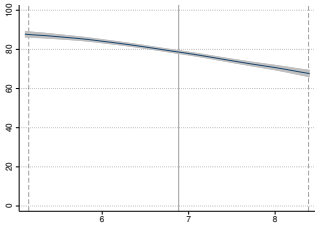
(r) Morocco



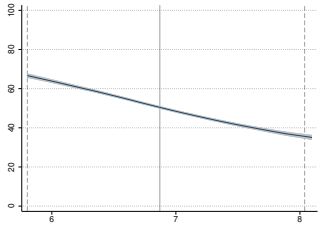
(s) Mozambique



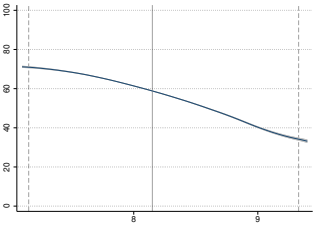
(t) Niger



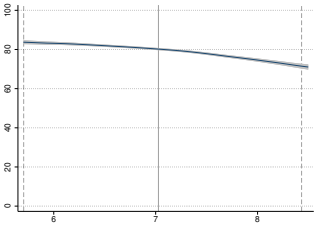
(u) Papua New Guinea



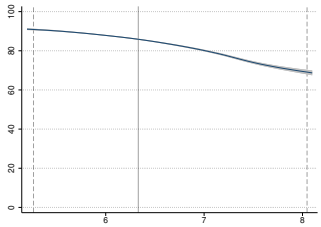
(v) Paraguay



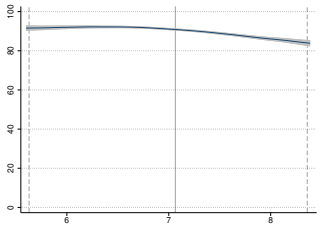
(w) Peru



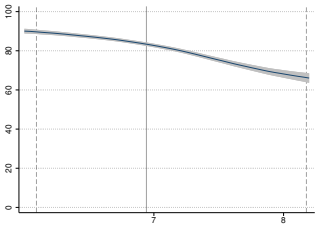
(x) Rep. of Congo



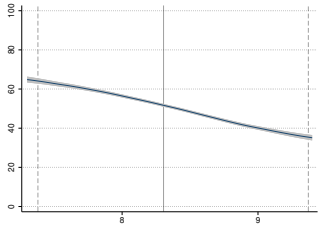
(y) Rwanda



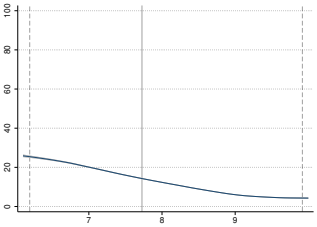
(z) Sao Tome



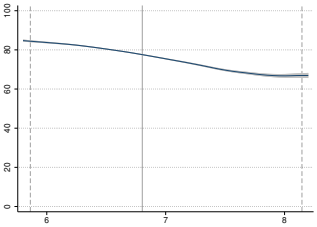
(aa) Senegal



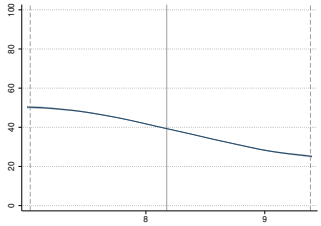
(ab) Serbia



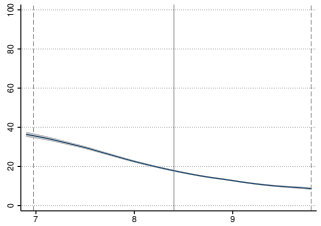
(ac) South Africa



(ad) Tanzania



(ae) Tunisia



(af) Uruguay

Notes: This fiure plots local polynomial fits of the Informality Engel Curves in all 32 core sample countries. Per person total expenditure on the horizontal axis is measured in log. The informal budget share is on the vertical axis. 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. The construction of informality Engel curves is presented in section 4.1.

Figure E2: Average Share of Unspecified Category by COICOP

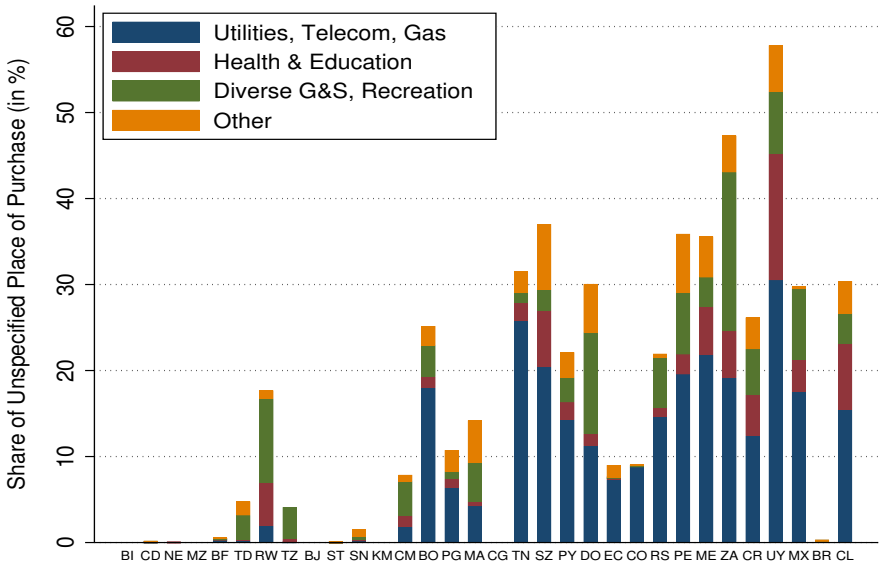
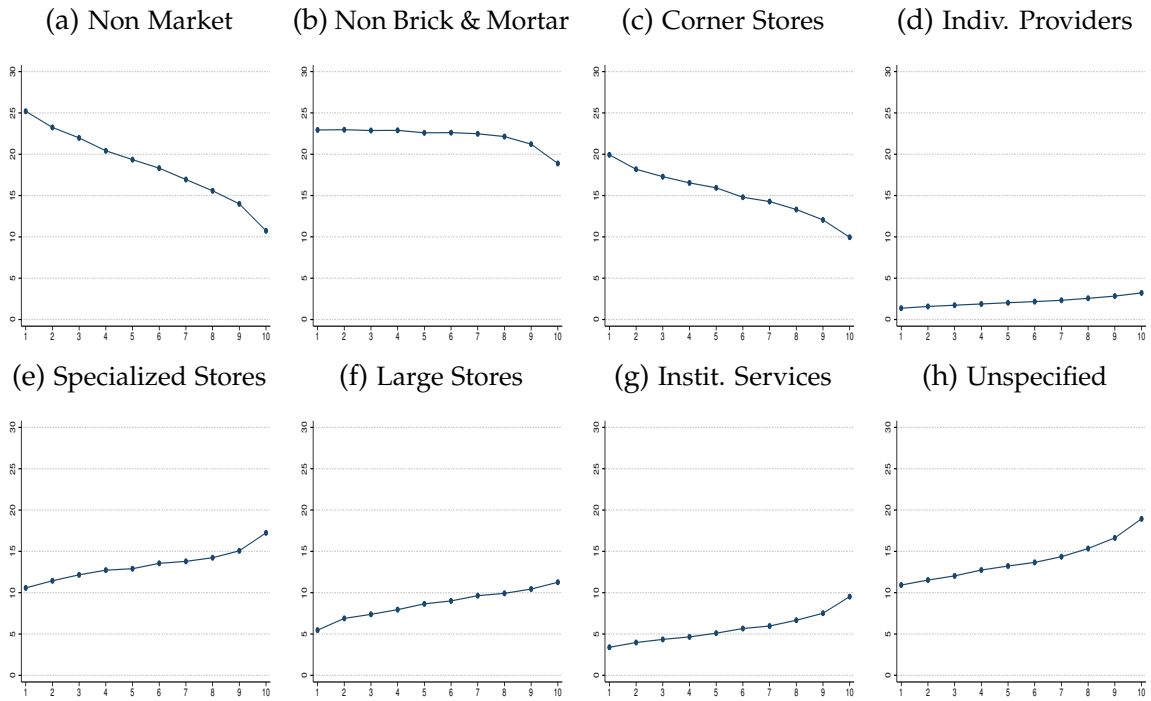


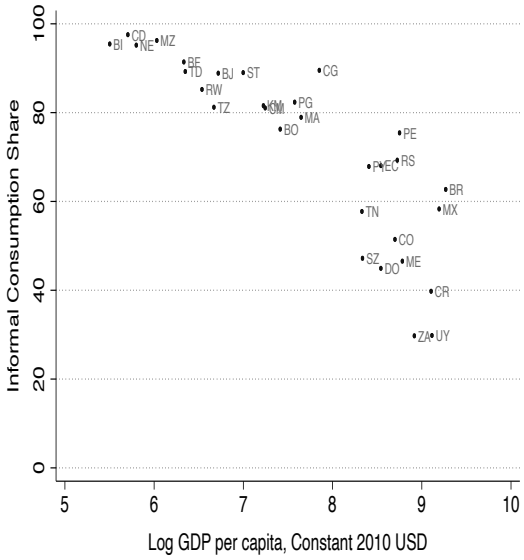
Figure E3: Average Expenditure of Each Decile by Place of Purchase



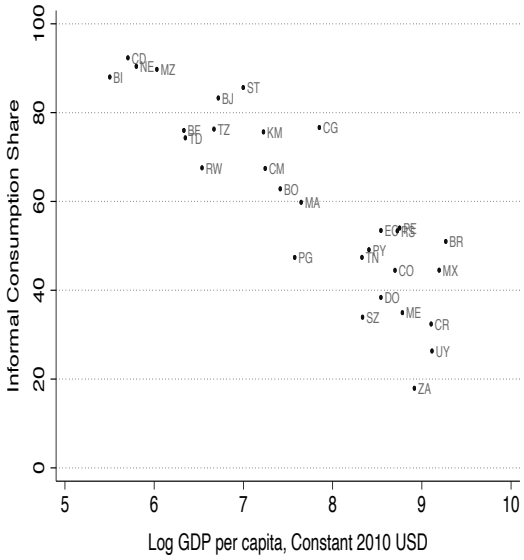
Notes: This figure shows the average cross-country expenditure by decile for different retailer-types, following the retailer taxonomy described in section 3.2. Panels (a), (b), (c), (d) show the places of purchase classified as informal and panels (e), (f), (g) and (h) show the places of purchase classified as formal in the central scenario of the paper.

Figure E4: Rural vs Urban Informal Consumption

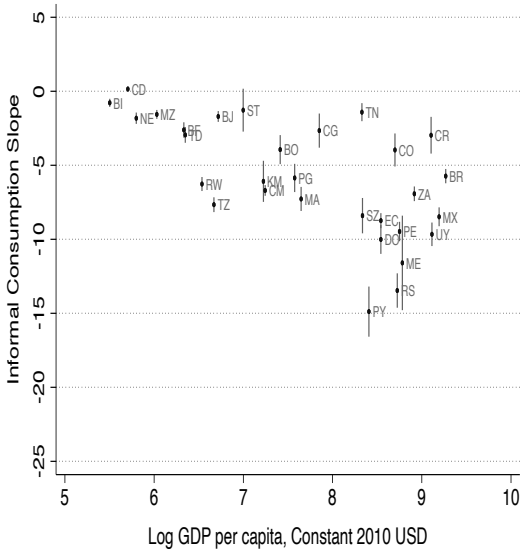
(a) Rural: Budget Share



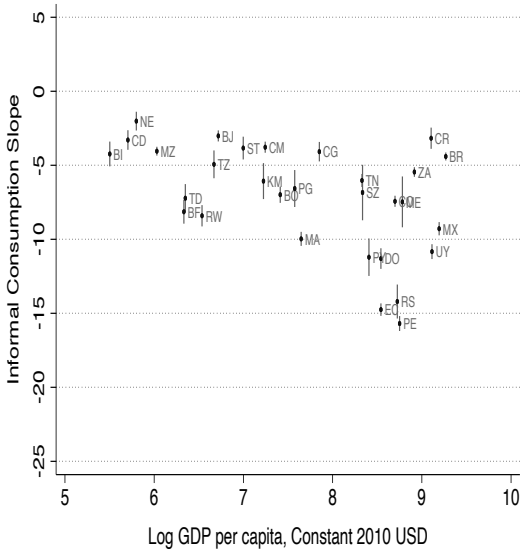
(b) Urban: Budget Share



(c) Rural: IEC slope



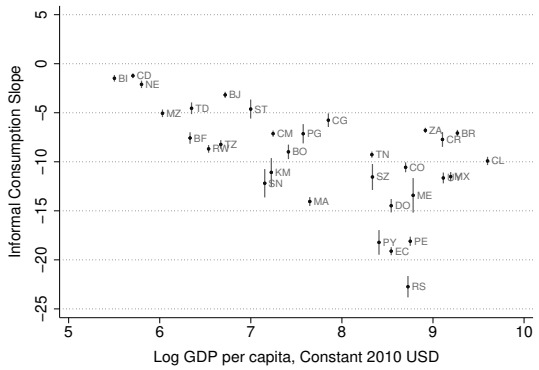
(d) Urban: IEC slope



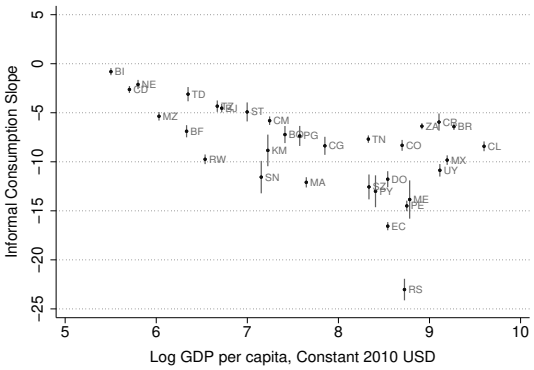
Notes: This figure plots informality levels and the slopes of the informality Engel curves for households located in rural regions (graphs a and c) and urban regions (graphs b and d). It only contains 30 out of our 32 sample countries, since the expenditure surveys in Chile and Senegal only contain urban populations.

Figure E5: Informality Engel Curve Slopes with Controls

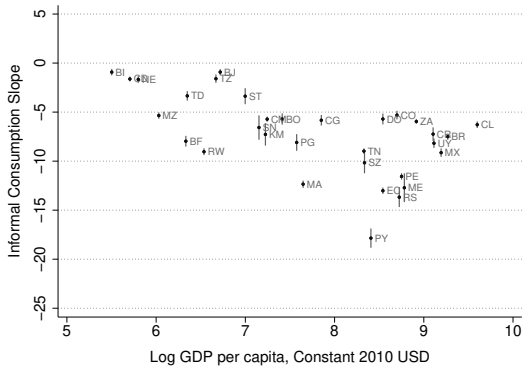
(a) Control: Rural Location



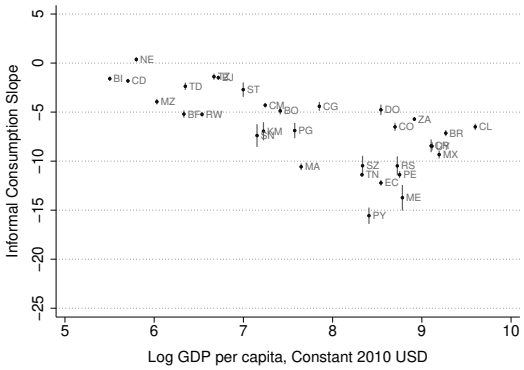
(b) Control: Survey Block



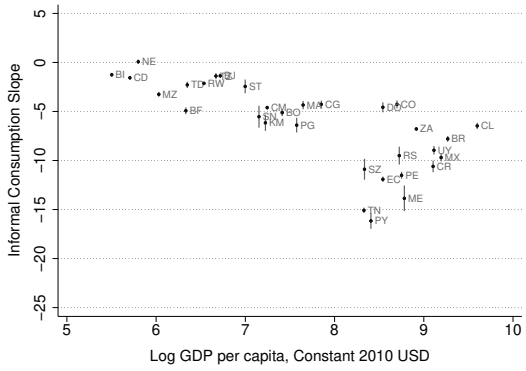
(c) Control: Food Goods



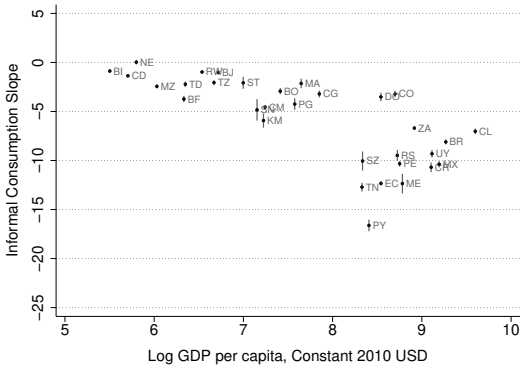
(d) Control: COICOP2 Goods



(e) Control: COICOP3 Goods

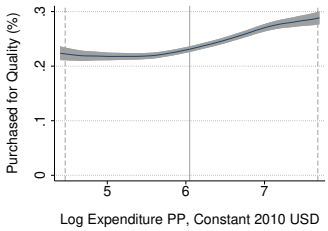


(f) Control: COICOP4 Goods

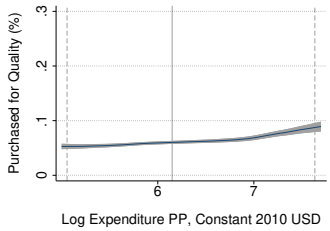


Notes: This figure shows countries' Informality Engel curve slopes against GDP per capita, controlling for geographical variables or increasingly narrow products. Panel (a) controls for an indicator if the household resides in a rural location while panel (b) controls for survey enumeration blocks. Panel (c) controls for food products, panel (d) for the 12 COICOP2 good categories, panel (e) for the 47 COICOP3 categories, and panel (f) for the 117 COICOP4 categories. More details in Section 4.2.

Figure E6: Share of Purchases where Store is Chosen for its Quality by Income

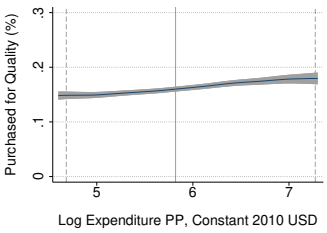


(a) Benin

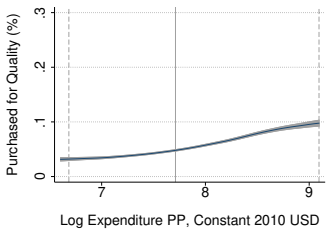


(b) Burundi

(c) Comoros



(d) Dem. Rep. of Congo



(e) Morocco

(f) Rep. of Congo

Notes: This figure shows local polynomial fits of the share of households buying any product for its quality against households' total expenditure per person (in log). Each panel corresponds to one of the six sample countries in which the expenditure survey asks respondents why they chose this place of purchase (for each expenditure). The solid vertical line corresponds to the median household total expenditure, while the dotted lines correspond to the 5th and 95th percentile. More details in Section 4.2.

Table E1: Household Expenditure Surveys

Country name	Survey	Year	Source	# HH	# items/HH	Exp/HH	Urban	HH Size	# PoP	# Modules	Product Code	Comments
						Cst. 2010 USD						
Benin	EMICOV	2015	World Bank	19872	31.9	261	48.2%	4.3	12	22	COICOP	
Bolivia	ECH	2004	Stat. Office	9149	49.4	585	60.7%	4.2	24	3	COICOP	
Brazil	POF	2009	Stat. Office	56049	48	3892	84.4%	3.3	33	8	Country-specific	
Burkina Faso	EICVM	2009	Stat. Office	8404	161.6	563	29.3%	6.7	45	1	COICOP	
Burundi	ECVM	2014	World Bank	6681	90.2	242	9.0%	4.8	13	23	COICOP	
Cameroon	ECAM	2014	World Bank	10303	95.8	1889	44.5%	4.6	17	1	COICOP	
Chad	ECOSIT	2003	World Bank	6747	92	356	10.9%	5.9	17	18	Country-specific	
Chile	EPF	2017	Stat. Office	15237	129.2	6872	100.0%	3.3	22	1	COICOP	No self-production, Only urban
Colombia	ENIG	2007	Stat. Office	42733	79.6	1850	82.4%	3.8	24	5	COICOP	
Comoros	EDMC	2013	Stat. Office	3139	83.5	1809	49.1%	5	12	19	COICOP	
Congo DRC	E123	2005	World Bank	12098	106.9	198	16.0%	5.3	13	1	COICOP	
Congo Rep	ECOM	2005	World Bank	5002	84.8	641	63.8%	5.1	17	1	COICOP	
Costa Rica	ENIGH	2014	Stat. Office	5705	67.5	5256	73.2%	3.4	41	1	COICOP	
Dominican Rep	ENIGH	2007	Stat. Office	8363	89.1	2396	67.6%	3.7	88	3	COICOP	
Ecuador	ENIGHUR	2012	World Bank	39617	88.6	1923	68.0%	3.9	75	7	COICOP	
Eswatini	HIES	2010	World Bank	3167	43.9	1283	37.4%	4.5	13	2	COICOP	
Mexico	ENIGH	2014	Stat. Office	19459	57.4	2272	64.5%	3.8	19	1	COICOP	
Montenegro	HBS	2009	World Bank	1223	148.9	3731	62.7%	3	7	3	COICOP	Cant separate categories 3 & 4
Morocco	ENCDM	2001	World Bank	14243	87.5	1679	61.6%	5.9	47	17	COICOP	
Mozambique	IOF	2009	World Bank	10809	48.7	363	28.9%	4.7	6	6	COICOP	
Niger	ENCBM	2007	World Bank	4000	221.2	325	17.2%	6.4	15	6	COICOP	
Papua NG	HIES	2010	World Bank	3811	111.2	1002	11.3%	5.1	6	1	COICOP	
Paraguay	EIGCV	2011	Stat. Office	5417	87,9	3466	61.2%	3.9	54	1	Country-specific	
Peru	ENAHO	2017	Stat. Office	43545	78.5	2609	76.8%	3.9	41	8	Country-specific	
Rwanda	EICV	2014	World Bank	14419	53.6	417	17.1%	4.6	11	8	COICOP	Pre-filled items
SaoTome	IOF	2010	World Bank	3145	105.9	705	68.1%	3.8	21	3	COICOP	
Senegal	EDMC	2008	World Bank	1443	517.8	640	100.0%	7.7	41	1	COICOP	Only urban
Serbia	HBS	2015	World Bank	6531	106	1888	61.9%	2.8	9	2	COICOP	
South Africa	IES	2011	U. of Cape Town	25325	44.2	3557	67.3%	3.8	6	1	COICOP	Cant separate categories 3 & 4
Tanzania	HBS	2012	World Bank	10186	317.8	478	21.9%	5	13	2	COICOP	Cant separate categories 3 & 4
Tunisia	ENBCNV	2010	Stat. Office	11281	139.1	1732	67.6%	4.3	9	1	COICOP	Cant separate categories 3 & 4
Uruguay	ENIGH	2005	Stat. Office	7042	77.5	2855	84.9%	3	39	1	COICOP	

Notes: The column '# PoP' refers to the number of different places of purchase in the country classification.

Table E2: IEC Slopes: Country Specific Formality Assignment

Specification:	**Main**		**Geography**		**Product Codes**				**All**
Avg. of 32 Countries	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Slope	9.2	9.9	8.5	7.6	6.4	5.9	5.7	5.2	3.9
Confidence Interval	[8.7,9.8]	[9.3,10.5]	[7.9,9.1]	[6.9,8.2]	[5.9,6.9]	[5.4,6.3]	[5.3,6.1]	[4.8,5.5]	[3.4,4.3]
# of p-values < 0.05	32	32	32	32	31	31	31	31	29
R^2 adjusted	0.20	0.22	0.28	0.44	0.45	0.53	0.53	0.53	0.57
Household Characteristics		X	X	X	X	X	X	X	X
Urban/Rural			X						
Survey Blocks				X					X
Food Products					X				
COICOP 2-dig						X			
COICOP 3-dig							X		
COICOP 4-dig								X	X

Table E3: IEC Slopes: Formality Assignment Based on Mexican VAT Reform

Specification:	**Main**		**Geography**		**Product Codes**				**All**
Avg. of 32 Countries	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Slope	9.6	10.4	8.9	7.9	6.9	6.4	6.2	5.7	4.3
Confidence Interval	[9.1,10.2]	[9.8,11.0]	[8.3,9.5]	[7.2,8.6]	[6.4,7.5]	[5.9,6.9]	[5.7,6.6]	[5.3,6.1]	[3.8,4.7]
# of p-values < 0.05	32	32	32	32	32	32	32	32	30
R^2 adjusted	0.20	0.23	0.29	0.44	0.44	0.51	0.52	0.51	0.55
Household Characteristics		X	X	X	X	X	X	X	X
Urban/Rural			X						
Survey Blocks				X					X
Food Products					X				
COICOP 2-dig						X			
COICOP 3-dig							X		
COICOP 4-dig								X	X

Notes: These tables show the (negative) average slope of the Informality Engel Curves across countries. Table E2 uses the country specific formality probabilities of modern and traditional stores. Table E3 assumes a 14% (77%) pass-through of taxes to informal (formal) prices, based on the Mexican reform estimates. Column 1 reports the slopes estimated from the following regression: $Share\ Informal_{ip} = \beta_0 + \beta_1 ln(expenditure_i) + \varepsilon_{ip}$ where $Share\ Informal_{ip}$ is the share of household i's informal expenditure on product p. Each observation is weighted using household survey weights and the expenditure share of the product. Average of lower and upper bounds of 95% confidence intervals are in brackets, calculated using robust standard errors. Column 2 augments this regression with controls for household characteristics (household size, age, gender, education of head). Column 3 (4) adds fixed effects for urban/rural (survey enumeration blocks). Column 5 instead adds fixed effects for food versus non-food products. Columns 6/7/8 instead add fixed effects for product codes at 2nd/3rd/4th level of the COICOP classification. Column 9 adds household characteristics and fixed effects for survey blocks and COICOP-4.

Table E4: IEC Slopes by Country

Country	Main		Geography			Product Codes			All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Benin	3.31	3.61	3.18	4.54	0.92	1.49	1.36	1.03	1.26
	(0.15)	(0.16)	(0.16)	(0.22)	(0.16)	(0.11)	(0.10)	(0.10)	(0.15)
Bolivia	9.77	11.43	8.99	7.22	5.71	4.87	5.13	2.93	2.74
	(0.29)	(0.33)	(0.38)	(0.44)	(0.29)	(0.19)	(0.18)	(0.16)	(0.25)
Brazil	7.60	7.98	7.07	6.41	7.50	7.15	7.79	8.11	6.64
	(0.15)	(0.17)	(0.17)	(0.18)	(0.16)	(0.16)	(0.15)	(0.13)	(0.14)
Burkina Faso	9.71	10.56	7.58	6.89	7.97	5.20	4.92	3.73	2.39
	(0.30)	(0.32)	(0.30)	(0.32)	(0.28)	(0.19)	(0.18)	(0.17)	(0.19)
Burundi	2.00	2.47	1.48	0.81	0.93	1.59	1.26	0.88	0.33
	(0.16)	(0.17)	(0.17)	(0.18)	(0.17)	(0.12)	(0.10)	(0.10)	(0.12)
Cameroon	8.21	9.35	7.13	5.81	5.72	4.30	4.61	4.55	2.88
	(0.13)	(0.14)	(0.16)	(0.22)	(0.13)	(0.12)	(0.10)	(0.09)	(0.13)
Chad	5.72	6.21	4.54	3.10	3.35	2.37	2.29	2.23	0.90
	(0.29)	(0.30)	(0.30)	(0.37)	(0.25)	(0.19)	(0.16)	(0.15)	(0.22)
Chile	9.92	9.91	9.91	8.42	6.28	6.50	6.47	7.02	5.97
	(0.20)	(0.21)	(0.21)	(0.25)	(0.17)	(0.17)	(0.16)	(0.15)	(0.17)
Colombia	9.76	10.52	10.56	8.32	5.31	6.51	4.28	3.22	3.37
	(0.23)	(0.25)	(0.26)	(0.28)	(0.22)	(0.21)	(0.20)	(0.17)	(0.19)
Comoros	9.54	11.65	11.08	8.84	7.28	6.95	6.16	5.93	4.42
	(0.58)	(0.71)	(0.74)	(0.82)	(0.58)	(0.47)	(0.42)	(0.37)	(0.56)
CongoDRC	1.35	2.22	1.23	2.63	1.62	1.82	1.57	1.36	1.39
	(0.13)	(0.15)	(0.13)	(0.18)	(0.13)	(0.11)	(0.10)	(0.08)	(0.14)
Congo Rep	6.38	7.50	5.76	8.37	5.83	4.41	4.27	3.21	3.10
	(0.34)	(0.36)	(0.36)	(0.47)	(0.28)	(0.23)	(0.20)	(0.18)	(0.27)
Costa Rica	7.22	8.60	7.72	5.95	7.25	8.44	10.60	10.69	8.84
	(0.35)	(0.37)	(0.38)	(0.45)	(0.35)	(0.33)	(0.30)	(0.25)	(0.30)
Dominican Rep	14.39	14.89	14.48	11.78	5.70	4.76	4.57	3.52	2.36
	(0.31)	(0.35)	(0.35)	(0.42)	(0.28)	(0.27)	(0.26)	(0.23)	(0.25)
Ecuador	19.11	20.90	19.11	16.57	13.02	12.22	11.92	12.34	9.46
	(0.18)	(0.19)	(0.21)	(0.21)	(0.16)	(0.15)	(0.14)	(0.12)	(0.13)
Eswatini	11.64	12.38	11.55	12.56	10.17	10.47	10.89	10.05	9.88
	(0.51)	(0.62)	(0.67)	(0.65)	(0.55)	(0.51)	(0.54)	(0.50)	(0.51)
Mexico	12.01	13.57	11.51	9.83	9.14	9.33	9.70	10.39	7.09
	(0.20)	(0.23)	(0.24)	(0.25)	(0.22)	(0.20)	(0.20)	(0.16)	(0.19)
Montenegro	15.87	16.64	13.42	13.85	12.71	13.73	13.86	12.35	10.20
	(0.79)	(0.89)	(0.90)	(1.00)	(0.75)	(0.66)	(0.67)	(0.51)	(0.56)
Morocco	16.85	18.11	14.05	12.09	12.35	10.57	4.34	2.14	0.00
	(0.21)	(0.22)	(0.23)	(0.27)	(0.19)	(0.18)	(0.21)	(0.25)	(0.28)
Mozambique	5.46	6.27	5.05	5.36	5.35	3.94	3.26	2.44	2.07
	(0.19)	(0.20)	(0.20)	(0.22)	(0.16)	(0.14)	(0.13)	(0.11)	(0.14)
Niger	2.14	2.56	2.10	2.12	1.67	0.00	0.00	0.00	0.00
	(0.18)	(0.19)	(0.19)	(0.24)	(0.17)	(0.12)	(0.11)	(0.12)	(0.14)
Papua New Guinea	8.59	9.35	7.14	7.36	8.10	6.88	6.40	4.24	3.06
	(0.49)	(0.49)	(0.50)	(0.52)	(0.43)	(0.40)	(0.38)	(0.30)	(0.32)
Paraguay	20.06	22.02	18.22	13.01	17.85	15.56	16.16	16.62	9.60
	(0.53)	(0.54)	(0.64)	(0.83)	(0.50)	(0.43)	(0.41)	(0.30)	(0.43)
Peru	19.58	21.83	18.10	14.50	11.56	11.38	11.51	10.34	6.05
	(0.22)	(0.23)	(0.24)	(0.27)	(0.16)	(0.18)	(0.18)	(0.15)	(0.17)
Rwanda	9.90	10.61	8.68	9.75	9.04	5.23	2.14	0.97	0.09
	(0.19)	(0.20)	(0.20)	(0.25)	(0.18)	(0.12)	(0.08)	(0.08)	(0.09)
Sao Tome	4.07	4.53	4.62	4.92	3.38	2.71	2.45	2.08	2.49
	(0.42)	(0.50)	(0.50)	(0.49)	(0.42)	(0.37)	(0.36)	(0.31)	(0.32)
Senegal	15.20	12.19	12.19	11.56	6.57	7.39	5.53	4.83	4.47
	(0.67)	(0.74)	(0.74)	(0.84)	(0.63)	(0.59)	(0.57)	(0.56)	(0.65)
Serbia	20.91	24.24	22.74	23.03	13.67	10.48	9.50	9.48	8.47
	(0.58)	(0.58)	(0.56)	(0.56)	(0.51)	(0.49)	(0.47)	(0.29)	(0.29)
South Africa	6.52	7.60	6.80	6.37	5.96	5.72	6.78	6.70	6.13
	(0.11)	(0.13)	(0.14)	(0.16)	(0.12)	(0.11)	(0.09)	(0.08)	(0.10)
Tanzania	9.75	8.64	8.24	4.33	1.59	1.38	1.39	2.06	1.13
	(0.22)	(0.22)	(0.23)	(0.30)	(0.22)	(0.16)	(0.15)	(0.13)	(0.17)
Tunisia	12.00	11.50	9.27	7.69	8.98	11.38	15.08	12.71	11.06
	(0.14)	(0.15)	(0.16)	(0.19)	(0.14)	(0.11)	(0.14)	(0.23)	(0.26)
Uruguay	11.57	11.73	11.65	10.87	8.18	8.48	8.96	9.31	8.36
	(0.25)	(0.27)	(0.28)	(0.32)	(0.24)	(0.22)	(0.22)	(0.19)	(0.21)
All Countries (Mean)	9.8	10.6	9.2	8.5	6.9	6.3	6.1	5.4	4.3
Household Characteristics		X	X	X	X	X	X	X	X
Urban/Rural			X						
Survey Blocks				X					X
Food Products					X			X	
COICOP 2-dig						X			
COICOP 3-dig							X		
COICOP 4-dig								X	X

Notes: 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. Controls include household characteristics (household size, age, gender, and education of head), geographic indicators (urban/rural and survey enumeration blocks), and product codes for food vs non-food as well as the 2nd, 3rd and 4th level of the United Nation's COICOP classification.

Table E5: Main Reason for Choosing a Place of Purchase

	Outcome: Share of purchases (in %)								
	Benin			**Burundi**			**Comoros**		
Reason	Informal	Formal	All Stores	Informal	Formal	All Stores	Informal	Formal	All Stores
Access	39.3	29.9	39.0	49.9	41.5	49.8	38.6	16.4	36.2
Price	26.8	11.6	26.4	27.8	14.8	27.6	31.7	26.1	31.1
Quality	23.5	51.4	24.3	5.7	41.0	6.4	9.0	39.8	12.4
Store Attributes	7.6	3.3	7.4	3.8	0.8	3.7	14.3	6.0	13.4
Other	2.9	3.9	2.9	12.8	1.9	12.6	6.4	11.7	7.0

	Dem. Rep of Congo			**Morocco**			**Rep. of Congo**		
Reason	Informal	Formal	All Stores	Informal	Formal	All Stores	Informal	Formal	All Stores
Access	28.9	16.1	28.7	58.7	57.3	58.5	37.5	26.8	36.8
Price	34.4	27.2	34.3	22.5	6.4	20.1	33.3	20.0	32.4
Quality	16.3	46.5	16.6	3.9	19.7	6.3	12.2	45.0	14.3
Store Attributes	7.8	7.6	7.8	0.6	7.7	1.7	7.4	4.3	7.2
Other	12.7	2.7	12.6	14.3	8.9	13.5	9.7	3.8	9.3

Notes: This table reports the frequencies across all purchases by reason of choosing a place of purchase, and shows the average for the six countries in the core sample which ask this question: Benin, Burundi, Comoros, Congo Rep., Morocco and RD Congo. In all surveys seven reasons are listed which we classify into five categories as follows: access is defined as "The retailer is closer or more convenient" and "The good or service cannot be found elsewhere", price as "The good or services are cheaper", quality as "The goods or services are of better quality", store attributes as "The retailer offers credit" and "The retailer is welcoming or is a friend" and other as "Others reasons". Note that Morocco has a few additional small categories, which pertain to attributes of retailer. The table lists the frequency for all purchases of goods and excludes services, which are less comparable along these dimensions, although their inclusion does not impact the results.

Table E6: Traditional & Modern Retailers Average Size and VAT Exemption Thresholds

Country	Trad Store Sales (Mill. USD)	Modern Store Sales (Mill. USD)	VAT Threshold (Mill. USD)	Ratio Traditional (/VAT)	Ratio Modern (/VAT)
Benin	0.06	2.61	0.09	0.76	30.57
Bolivia	0.03	5.63	0.03	1.19	204.84
Brazil	0.06	2.83	0.61	0.10	4.65
Burkina Faso	0.11	3.12	0.09	1.32	36.60
Burundi	0.04	0.75	0.05	0.71	13.87
Cameroon	0.06	2.51	0.09	0.76	29.43
Chad	0.13	2.57	0.17	0.76	15.08
Chile	0.05	7.11	0.03	1.64	227.28
Colombia	0.05	2.31	0.81	0.06	2.85
Comoros	0.10	0.47	0.05	0.76	15.08
Congo. DRC	0.04	0.92	0.05	0.90	19.01
Congo. Rep	0.09	4.39	0.10	0.87	42.84
Costa Rica	0.07	4.99	0.11	0.65	45.12
Dominican Republic	0.06	7.49	0.16	0.38	47.58
Ecuador	0.04	6.58	0.06	0.64	109.64
Eswatini	0.07	0.70	0.03	1.95	20.63
Mexico	0.04	1.93	0.10	0.35	18.55
Montenegro	0.12	1.04	0.20	0.57	5.16
Morocco	0.07	1.40	0.21	0.34	6.72
Mozambique	0.17	2.54	0.08	2.08	31.71
Niger	0.12	2.48	0.10	1.14	24.53
Papua NG	0.06	1.24	0.07	0.75	16.76
Paraguay	0.06	1.24	0.07	0.75	16.76
Peru	0.03	2.52	0.11	0.27	23.35
Rwanda	0.06	1.04	0.02	2.51	46.70
Sao Tome	0.06	0.10	0.02	2.04	4.17
Senegal	0.10	4.46	0.10	0.97	44.11
Serbia	0.11	1.19	0.08	1.45	15.64
South Africa	0.14	2.80	0.10	1.39	27.29
Tanzania	0.05	0.12	0.04	1.11	2.73
Tunisia	0.07	1.95	0.05	1.26	34.74
Uruguay	0.13	3.10	0.08	1.53	37.02
Mean	**0.08**	**2.67**	**0.13**	**1.01**	**38.85**

Notes: This table reports the average sales of traditional and modern stores in the 32 countries in our sample, based on Euromonitor reports. In addition, it reports the value of the VAT exemption threshold, based on country-reports produced by the International Bureau of Fiscal Documentation. The final two columns report the ratio of the average traditional and modern store sales-value to the VAT exemption threshold.

Table E7: Country-Specific Places of Purchase

Assigned	%	Original name	Classification
		BENIN	
Formal	4.8	autre lieu d'achat formel sur le territoir	5 large stores
	0.6	achat dans un supermarche	5 large stores
	0.9	achat dans un magasin ou un atelier formel	4 specialized shops
	0.7	achat dans un magasin ou un atelier formel	4 specialized shops
	9.4	achat au secteur public ou parapublic	6 institutions
Informal	20.1	achat au domicile du vendeur, dans une pet	3 corner shops
	21.0	achat sur un marche public	2 no store front
	19.2	achat chez un ambulant, ou poste fixe sur	2 no store front
	14.2	bien ou service autoproduit	1 non-market
	8.2	autre lieu d'achat informel (independant)	1 non-market
		BOLIVIA	
Formal	0.9	supermercado	5 large stores
	11.8	tienda especializada	4 specialized shops
	3.6	instituto educativo	6 institutions
	1.3	institucionn de salud	6 institutions
	0.6	comunicacion	6 institutions
	1.5	hotel, bar, restaurante	8 entertainment
Informal	14.7	tienda de conveniencia	3 corner shops
	19.7	mercado	2 no store front
	4.0	feria	2 no store front
	3.6	puesto/kiosco	2 no store front
	2.1	vendedor ambulante	2 no store front
	1.9	de un hogar / transferencia	1 non-market
	1.5	auto consumo	1 non-market
	5.6	cantina	9 informal entertainment
Unspec.	25.8	other	99 n.a./other
		BRAZIL	
Formal	14.0	supermarket	5 large stores
	1.0	department store	5 large stores
	28.9	specialized shop	4 specialized shops
	8.4	vehicle	4 specialized shops
	5.2	pharmacy	4 specialized shops
	3.3	education institution	6 institutions
	2.0	health institution	6 institutions
	1.0	bank	6 institutions
	0.9	public health	6 institutions
	0.8	internet	6 institutions
	0.6	lottery	6 institutions
	0.5	communication company	6 institutions
	3.4	restaurant	8 entertainment
Informal	4,5	grocery store	3 corner shops
	1.8	street seller	2 no store front
	1.7	small market	2 no store front
	1.4	fair	2 no store front
	1.0	small shop	2 no store front
	0.7	own production_from other household	1 non-market
	0.5	from farm	1 non-market
	13.7	private service	7 service from individual
	2.6	bar-cafe	9 informal entertainment
	0.6	recreation events	9 informal entertainment
Unspec.	0,1	other	99 n.a./other
		BURKINA FASO	
Formal	1.1	grands magasin	5 large stores
	0.9	magasin de gros a petits prix	5 large stores
	3.8	station service (lubrifiants)	4 specialized shops
	1.8	pharmacie	4 specialized shops
	1.6	atelier, service reparation	4 specialized shops
	3.8	ecole, lycees, universite privas	6 institutions
	2.0	ecole, lycee, universite publics	6 institutions
	1.2	clinique, laboratoire medical public	6 institutions
	0.9	telephone, eau, electricite	6 institutions
	0.7	cabine telephone privee	6 institutions
	1.4	bar, cafe, restaurant, hotel	8 entertainment
Informal	12.4	boutique de quartier	3 corner shops
	1.2	quincallerie (petite taille)	3 corner shops
	38.6	marche	2 no store front
	1.1	marchant ambulants	2 no store front
	0.7	kiosque ou echoppe quartier	2 no store front
	11.0	menage	1 non-market
	6.7	bien ou service autoproduit	1 non-market
	1.5	cadeau recu en nature ou en espace	1 non-market
	2.5	autres service prives	7 service from individual
	1.3	service de transport prive	7 service from individual
Unspec.	0.4	other	99 n.a./other
		BURUNDI	
Formal	5.5	autre lieu d'achat formel	5 large stores
	1.4	magasin, atelier formel (societe) tenu	4 specialized shops
	0.7	magasin ou atelier formel (societe)	4 specialized shops
	3.3	secteur public ou parapublic	6 institutions
	0.7	hors lieu de residence ou a l'etranger	6 institutions
Informal	17.7	autre lieu d'achat informel	3 corner shops
	29.6	marche public	2 no store front
	24.7	domicile du vendeur, petite boutique	2 no store front
	4.7	vendeur ambulant ou poste fixe sur voie	2 no store front
	10.5	bien ou service autoproduit	1 non-market
	1.1	cadeau recu	1 non-market

Assigned	%	Original name	Classification
		CAMEROON	
Formal	0.9	supermarche/grand magasin	5 large stores
	4.0	magasin specialistes	4 specialized shops
	6.8	secteur transport	6 institutions
	3.2	presetation de services publics	6 institutions
	3.0	cliniques	6 institutions
	5.1	hotels/bars/restaurants	8 entertainment
Informal	8.8	epiceries/boutiques/echoppes	3 corner shops
	22.4	marches	2 no store front
	3.2	kiosque de jeux et call box	2 no store front
	3.1	vente ambulante	2 no store front
	0.8	vendeurs specialises hors magasins	2 no store front
	15.2	don. cadeau recu	1 non-market
	7.7	auto production	1 non-market
	2.9	dans la nature/forit/brousse	1 non-market
	2.3	domicile de vendeur	1 non-market
	2.8	prestation de services individuels	7 service from individual
Unspec.	7.7	other	99 n.a./other
		CHAD	
Formal	0.7	supermarche	5 large stores
	5.7	boutique	4 specialized shops
	1.1	magasins	4 specialized shops
	4.6	autre prestataire de service privee	6 institutions
	1.0	autre prestataire de service public	6 institutions
	0.7	transport privee	6 institutions
	0.7	enseignement privee	6 institutions
	0.6	prestataire service sante privee	6 institutions
	0.6	enseignement public	6 institutions
	0.5	prestataire service sante public	6 institutions
	1.6	hotel. restaurant. ..	8 entertainment
Informal	0.4	echoppe	3 corner shops
	30.4	marche centraux	2 no store front
	25.6	marche de quartier ou specialiste	2 no store front
	2.5	marchand ambulant	2 no store front
	1.1	tablier	2 no store front
	16.7	self-consumption	1 non-market
Unspec.	5.0	other	99 n.a./other
		CHILE	
Formal	13.3	supermercados	5 large stores
	3.3	multitiendas	5 large stores
	0.9	distribuidoras - mayoristas	5 large stores
	26.5	tienda especializada	4 specialized shops
	4.9	comercio ambulante	4 specialized shops
	2.4	farmacias	4 specialized shops
	1.6	ferreterias y multiferreterias	4 specialized shops
	2.0	clinicas	6 institutions
	1.0	extranjero	6 institutions
	4.3	restaurantes y bares	8 entertainment
Informal	5.4	almacen tradicional	3 corner shops
	0.5	otros establecimientos	3 corner shops
	2.9	ferias libres	2 no store front
Unspec.	30.3	other	99 n.a./other
		COLOMBIA	
Formal	33.0	almacenes o supermercados de cadena y tien	5 large stores
	9.9	hipermercados	5 large stores
	2.0	centrales mayoristas de abastecimiento	5 large stores
	0.8	plazas de mercado y galerias	5 large stores
	0.5	cooperativas. fondos de empleados y comisa	5 large stores
	10.7	farmacias y droguerias	4 specialized shops
	1.7	restaurantes	8 entertainment
Informal	13.3	miscelÍÁneas de barrio y cacharrerias	3 corner shops
	4.6	tiendas de barrio	3 corner shops
	1.7	sanandresitos	2 no store front
	1.4	ferias especializadas: artesanal. del hoga	2 no store front
	1.1	persona particular	1 non-market
	0.9	transfers. from household	1 non-market
	0.5	self production	1 non-market
	5.6	cafeterias y establecimientos de comidas	9 informal entertainment
Unspec.	10.5	other	99 n.a./other
		COMOROS	
Formal	11.1	autre lieu d'achat formel	5 large stores
	2.2	supermarche	5 large stores
	5.4	magasin. atelier formel (societe) tenu	4 specialized shops
	2.6	magasin ou atelier formel (societe)	4 specialized shops
	5.9	secteur public ou parapublic	6 institutions
	2.4	hors lieu de residence ou a l'etranger	6 institutions
Informal	31.6	domicile du vendeur. petite boutique	3 corner shops
	20.6	marche public	2 no store front
	6.5	vendeur ambulant ou poste fixe sur voie	2 no store front
	6.3	autre lieu d'achat informel	1 non-market
	3.6	bien ou service autoproduit	1 non-market
	1.8	cadeau recu	1 non-market
		CONGO DRC	
Formal	0.5	achat supermarche	5 large stores
	3.8	achat magasin non indo-pakistanais	4 specialized shops
	3.2	achat magasin indo-pakistanais	4 specialized shops
	3.1	achat secteur public	6 institutions
Informal	36.5	achat marche public	2 no store front
	10.1	achat ambulant	2 no store front
	5.8	autre lieu informel	2 no store front
	17.9	achat domicile	1 non-market
	17.5	bien ou service autoproduit	1 non-market
	1.4	cadeau recu	1 non-market
Unspec.	0.1	other	99 n.a./other

COSTA RICA

Assigned	%	Original name	Classification
Formal	17.0	supermercado	5 large stores
	1.3	tienda de departamentos	5 large stores
	11.3	local especializado	4 specialized shops
	4.2	gasolinera y estacion de servicio	4 specialized shops
	3.6	tienda de ropa, zapateria, perfumeria	4 specialized shops
	1.1	carniceria pescaderia	4 specialized shops
	1.0	salones de estetica o belleza	4 specialized shops
	1.9	laboratorio, clinica, centro medico	6 institutions
	1.1	en el exterior	6 institutions
	3.9	restaurante, soda, cafeteria, heladeria	8 entertainment
	3.4	almacen de electrodomisticos y de tecnol	8 entertainment
	1.7	comedor en lugar de trabajo	8 entertainment
	0.8	retiro del negocio	8 entertainment
Informal	6.2	pulperia o minisuper	3 corner shops
	2.4	vendedor ambulante o a domicilio	2 no store front
	0.8	local de articulos usados	2 no store front
	8.9	recibido o comprado a otros hogares	1 non-market
Unspec.	25.1	other	99 n.a./other

DOMINICAN REPUBLIC

Formal	3.9	tienda por departamentos	5 large stores
	3.6	supermercados	5 large stores
	2.3	farmacias	4 specialized shops
	1.4	salon de belleza	4 specialized shops
	1.2	tienda de ropa	4 specialized shops
	1.2	tienda de electrodomesticos	4 specialized shops
	1.1	taller de mecanica en general, desabulladu	4 specialized shops
	1.0	ferreterias	4 specialized shops
	0.8	carniceria	4 specialized shops
	0.7	tienda de respuestos de vehiculos	4 specialized shops
	0.6	puesto de rifa de aguante y loteria electr	4 specialized shops
	0.5	peluqueria	4 specialized shops
	3.5	estacionn de gasolina	6 institutions
	2.0	clinica	6 institutions
	1.9	envasadora de gas	6 institutions
	1.8	comedor popular	6 institutions
	1.7	corporacion de electricidad	6 institutions
	1.5	colegio	6 institutions
	1.3	hospitales	6 institutions
	1.2	compania de telefonos	6 institutions
Informal	20.5	colmado	3 corner shops
	0.7	almacen de provisiones	3 corner shops
	3.2	vendedora ambulante	2 no store front
	1.2	mercados	2 no store front
	1.0	puestos de venta	2 no store front
	0.6	picapollo	2 no store front
	1.9	autosuministro	1 non-market
	1.5	cafeteria	9 informal entertainment
Unspec.	29.5	other	99 n.a./other

ECUADOR

Formal	4.1	supermercados de cadena	5 large stores
	1.2	hipermercados	5 large stores
	4.2	ropa de todo tipo	4 specialized shops
	2.2	electrodomesticos y accesorios	4 specialized shops
	2.0	calzado de todo tipo	4 specialized shops
	1.8	otros sitios de compra especializados	4 specialized shops
	1.4	librerias y papelerias	4 specialized shops
	1.2	panaderas	4 specialized shops
	1.2	mecanicas automotrices	4 specialized shops
	1.1	gasolineras	4 specialized shops
	1.0	salas de belleza	4 specialized shops
	0.8	muebles y enceres	4 specialized shops
	0.6	tercena/carnicera	4 specialized shops
	0.5	repuestos de automotores	4 specialized shops
	5.2	boticas y farmacias	4 specialized shops
	5.0	establecimientos educativos	6 institutions
	2.3	transporte de pasajeros	6 institutions
	2.2	establecimientos privados de salud	6 institutions
	1.9	servicios profesionales (abogados. arqu)	6 institutions
	1.2	venta por catalogo o television	6 institutions
	0.9	bahas, ipiales	6 institutions
	0.6	instituciones publicas	6 institutions
	0.6	aseguradoras	6 institutions
	2.4	restaurantes, salones	8 entertainment
	0.5	centros. serv. de recreacion. estadios	8 entertainment
Informal	13.1	tiendas de barrio	3 corner shops
	1.6	bodegas, distribuidores	3 corner shops
	10.6	mercados	2 no store front
	2.1	vendedores ambulantes	2 no store front
	1.1	ferias libres	2 no store front
	11.4	productos autoconsumo. autosuministro	1 non-market
	1.0	personas particulares	7 service from individual
Unspec.	8.5	other	99 n.a./other

ESWATINI

Assigned	%	Original name	Classification
Formal	27.6	supermarket	5 large stores
	5.6	clothes/footwear/linen	4 specialized shops
	1.7	hardware store	4 specialized shops
	1.4	butchery	4 specialized shops
Informal	5.8	grocery	3 corner shops
	0.6	spaza	3 corner shops
	4.0	street vendor	2 no store front
	1.9	market	2 no store front
	7.0	self production	1 non-market
	5.5	gifts/transfers	1 non-market
Unspec.	38.8	other	99 n.a./other

MEXICO

Formal	11.5	supermercados	5 large stores
	2.1	tiendas departamentales	5 large stores
	1.0	tiendas con membresia	5 large stores
	21.3	tiendas especificas del ramo	4 specialized shops
	0.7	diconsa	6 institutions
	0.5	compras fuera del pais	6 institutions
	2.4	restaurantes	8 entertainment
Informal	12.9	tiendas de abarrotes	3 corner shops
	0.6	tiendas de conveniencia	3 corner shops
	5.7	persona particular	2 no store front
	3.7	mercado	2 no store front
	3.1	vendedores ambulantes	2 no store front
	2.0	tianguis o mercado sobre ruedas	2 no store front
	2.6	loncherias, fondas, torterias, cocina	9 informal entertainment
Unspec.	29.2	other	99 n.a./other

MONTENEGRO

Formal	17.2	supermarket	5 large stores
	36.2	store	4 specialized shops
Informal	5.3	stall	2 no store front
	5.3	own production	1 non-market
Unspec.	35.8	other	99 n.a./other

MOROCCO

Formal	0.7	supermarket or hypermarket	5 large stores
	5.2	butcher or retail chicken seller	4 specialized shops
	3.1	pharmacy	4 specialized shops
	1.9	craftsman s shop (hairdresser. tailor. etc	4 specialized shops
	1.7	shop for selling furniture and durable ite	4 specialized shops
	1.4	modern clothes shop	4 specialized shops
	1.0	gas stations (benzine. etc.)	4 specialized shops
	0.8	bookshop (small bookshop or kiosk in the n	4 specialized shops
	0.7	pastry shop. bakery or snack-bars	4 specialized shops
	0.5	retail fish seller	4 specialized shops
	6.4	public and semi public agencies	6 institutions
	2.1	regular transportation means (bus. train.	6 institutions
	2.1	medical care in a private institution	6 institutions
	1.2	public administration	6 institutions
	1.1	public baths. shower. swimming pool	6 institutions
	0.7	private education institution	6 institutions
	0.5	banks. financing institutions and insuranc	6 institutions
Informal	16.2	neighbourhood or village grocer	3 corner shops
	3.1	grocers	3 corner shops
	0.5	greengrocers	3 corner shops
	17.9	weekly market	2 no store front
	4.4	neighbourhood market	2 no store front
	2.0	itinerant merchant selling on sidewalks	2 no store front
	0.9	city market or central market	2 no store front
	4.7	self-production	1 non-market
	1.2	cafe. non standing restaurant	9 informal entertainment
Unspec.	13.8	other	99 n.a./other

MOZAMBIQUE

Formal	35.8	outro	4 specialized shops
	11.5	loja	4 specialized shops
Informal	16.5	mercado informal	2 no store front
	11.9	mercado	2 no store front
	24.0	auto produco	1 non-market

NIGER

Formal	0.1	supermarche/grand magasin	5 large stores
	3.1	secteur transport	6 institutions
	1.4	clinique. laboratoire. ecole	6 institutions
Informal	32.8	epicerie. boutique	3 corner shops
	19.2	marche	2 no store front
	4.8	vente ambulante	2 no store front
	17.6	auto production	1 non-market
	9.7	prestation services publiques	1 non-market
	3.8	cadeau recu	1 non-market
	6.6	prestation service individuels	7 service from individual

PAPUA NEW GUINEA

Assigned	%	Original name	Classification
Formal	34.5	supermarket	5 large stores
Informal	9.4	small shop. canteen. tuck shop	3 corner shops
	10.5	local market	2 no store front
	3.8	street vendor	2 no store front
	14.2	home production	1 non-market
	10.2	gift	1 non-market
Unspec.	17.6	other	99 n.a./other

PARAGUAY

Formal	13.2	supermercado	5 large stores
	4.4	estacion de servicio	4 specialized shops
	2.8	farmacia	4 specialized shops
	1.6	empresa de transporte	4 specialized shops
	1.5	carniceria	4 specialized shops
	0.7	muebleria	4 specialized shops
	0.6	joyerias	4 specialized shops
	0.6	peluqueria	4 specialized shops
	0.6	libreria	4 specialized shops
	2.8	cubierto por el seguro	6 institutions
	0.8	instituto	6 institutions
	1.4	restaurantes y o bares	8 entertainment
Informal	13.6	despensa	3 corner shops
	4.8	puesto fijo	3 corner shops
	2.4	tienda	3 corner shops
	2.6	ambulante	2 no store front
	1.9	mercado	2 no store front
	4.8	regalado o pagado por algien miembro de	1 non-market
	3.7	producido por el hogar	1 non-market
	2.4	regalado o donado por algien programa so	1 non-market
	2.2	retirado del negocio	1 non-market
	1.5	como parte de pago a un miembro del hogar	1 non-market
	3.0	consultorio privado	7 service from individual
Unspec.	22.4	other	99 n.a./other

PERU

Formal	2.0	supermercado	5 large stores
	0.6	bodega (x mayor)	5 large stores
	3.4	tienda especializada al por menor	4 specialized shops
	2.1	farmacia	4 specialized shops
	0.5	libreria	4 specialized shops
	2.0	empresas de transporte formales	6 institutions
	0.9	centro de estudios	6 institutions
	0.8	grifos de empresas	6 institutions
	0.6	restaurantes y/bares	8 entertainment
Informal	8.7	bodega (x menor)	3 corner shops
	14.2	mercado (x menor)	2 no store front
	2.9	ambulante	2 no store front
	2.0	mercado (x mayor)	2 no store front
	1.6	feria	2 no store front
	13.8	self-consumption	1 non-market
Unspec.	34.6	other	99 n.a./other

REPUBLIC OF CONGO

Formal	1.0	grands magasins	5 large stores
	7.0	autres commerces modernes	4 specialized shops
	5.8	prestataires de services publics	6 institutions
	3.9	secteur transports	6 institutions
	2.5	cliniques, laboratoires médicaux	6 institutions
	3.9	hotels. restaurants. bars. cafes	8 entertainment
Informal	3.4	epiceries modernes	3 corner shops
	42.8	marches	2 no store front
	8.4	echoppes sur marches et sur bord de route	2 no store front
	6.2	marchands ambulants	2 no store front
	4.5	produit autoconsommes	1 non-market
	3.9	ménages	1 non-market
	5.5	prestataires de services individuels	7 service from individual

RWANDA

Formal	0.5	supermarket/big shop	5 large stores
	3.8	specialized shop	4 specialized shops
	4.9	bar/restaurant	8 entertainment
Informal	18.4	small shop/boutique	3 corner shops
	12.0	market	2 no store front
	2.0	individual	2 no store front
	0.9	mobile seller	2 no store front
	25.7	self production	1 non-market
	3.0	from a household	1 non-market
	11.2	service provider	7 service from individual
Unspec.	17.8	other	99 n.a./other

SAO TOME

Formal	5.4	grandes lojas	5 large stores
	5.4	lojas modernas	5 large stores
	1.3	outros comercios modernos	4 specialized shops
	0.8	sector de transportes	6 institutions
	1.0	hotels. restaurantes. bares. cafes	8 entertainment
Informal	33.6	quiosque / quitanda	3 corner shops
	23.9	mercado	2 no store front
	7.8	vendedor ambulante	2 no store front
	5.9	agregados	1 non-market
	4.5	prestates de servicios publicos	1 non-market
	1.9	auto consumo	1 non-market
	0.9	campo. mato	1 non-market
	0.6	autoabastecimento	1 non-market
	3.8	prestates de servicios individuais	7 service from individual
	1.6	candongueiro	7 service from individual
Unspec.	0.1	other	99 n.a./other

SENEGAL

Assigned	%	Original name	Classification
Formal	0.4	magasins de gros ou a prix reduits	5 large stores
	2.6	station service (carburants. lubrifiants.e	4 specialized shops
	1.0	boulangerie, patisserie	4 specialized shops
	0.6	boucherie	4 specialized shops
	1.1	societe de telephonie et de distributi	6 institutions
	0.9	service de transport public	6 institutions
	0.5	ecole. lycee. universitee prives	6 institutions
	7.4	bar, cafe, restaurant, hotel	8 entertainment
Informal	34.7	boutique de quartier	3 corner shops
	28.5	marches	2 no store front
	6.9	kiosque ou échoppe au quartier	2 no store front
	1.0	marchand ambulant	2 no store front
	2.2	cadeau recu en nature	1 non-market
	1.2	bien ou service autoproduit	1 non-market
	4.6	service de transport privé	7 service from individual
	1.3	autres services prives	7 service from individual
Unspec.	1.7	other	99 n.a./other

SERBIE

Formal	8.9	hypermarket	5 large stores
	23.8	specialized shop	4 specialized shops
	2.9	discounted shop	4 specialized shops
Informal	29.6	minimarket	3 corner shops
	4.8	market/open	2 no store front
	1.8	gray economy	2 no store front
	5.3	own production/own business	1 non-market
	2.2	gifts/received transfers	1 non-market
Unspec.	20.7	other	99 n.a./other

SOUTH AFRICA

Formal	38.6	chain store	5 large stores
	11.2	other retailer	4 specialized shops
Informal	2.7	other	2 no store front
	0.9	street trading	2 no store front
	0.6	from a household	1 non-market
Unspec.	45.7	other	99 n.a./other

TANZANIA

Formal	0.8	duka kubwa(department stores)	5 large stores
	36.9	shop	4 specialized shops
Informal	15.8	market	2 no store front
	2.5	street vendor	2 no store front
	25.0	produced by household	1 non-market
	4.1	other household	1 non-market
	1.3	gift or free	1 non-market
Unspec.	5.5	other	99 n.a./other

TUNISIA

Formal	1.2	hyper, supermarche	5 large stores
	60.4	boutique privee	4 specialized shops
Informal	4.2	ambulant	2 no store front
	1.1	point de vente marche	2 no store front
	1.5	cadeau	1 non-market
	1.2	auto production	1 non-market
Unspec.	30.3	other	99 n.a./other

URUGUAY

Formal	11.7	autoservicio. cadena de supermercados	5 large stores
	1.0	shopping o galeria	5 large stores
	2.6	carniceria. polleria. pescaderia	4 specialized shops
	2.3	merceria. tienda	4 specialized shops
	1.5	panaderia. confiteria	4 specialized shops
	1.3	casa de electrodomesticos. telefonos	4 specialized shops
	0.9	zapateria. marroquineria. talabarteria	4 specialized shops
	0.7	farmacia. perfumeria. panalera	4 specialized shops
	0.7	verduleria. puesto. fruteria	4 specialized shops
	0.8	fuera del pais	6 institutions
	0.8	restaurante. parrillada	8 entertainment
	0.5	cantina. trabajo. colegio	8 entertainment
Informal	7.7	almacen	3 corner shops
	1.5	feria vecinal	2 no store front
	1.0	vendedor ambulante. puesto callejero. carr	2 no store front
	0.7	quiosco. salon	2 no store front
	0.8	bar. pizzeria	9 informal entertainment
Unspec.	59.9	other	99 n.a./other