Incidence of the Value Added Tax in the context of high informality

Ana P. Franco and Luciana Galeano*

Preliminary draft. Version: November 2024

Abstract

This paper presents an in-depth study on the incidence of the value-added tax (VAT) in high-informality settings. While consumption taxes have traditionally been considered regressive, recent research suggests they might be progressive under specific assumptions regarding pass-through of taxes to prices and household shopping behavior. We use high-frequency price data from formal and partially informal markets in Peru, taking advantage of a natural experiment in the form of a temporary VAT exemption to calculate pass-through in the formal and informal sector. We also use household survey data to analyze consumption patterns across the income distribution. Preliminary results suggest that households in the lower part of income distribution disproportionately shop at informal markets for their purchases (the bottom 10% of the income distribution spend about 34 percentage points more on informal markets than the top 10%). At the same time, pass-through rates of taxes to prices in formal markets significantly exceed those in informal markets, although the latter is not necessarily 0 as recent literature suggests. While pass-through in formal markets is close to 100%, pass-through in informal markets is in general lower and varies by product (from 0% to 100%). We also find that the informal markets exhibit anticipatory reactions prior to the policy's effective

*University of Michigan. apfranco@umich.edu; lgaleano@umich.edu. We are grateful to Joel Slemrod, James Hines Jr., Mazhar Waseem, Hoyt Bleakley, and participants in the Public Finance Seminar of the University of Michigan for their valuable insights and feedback. We are also grateful to the Ministry of Agricultural Development and Irrigitation in Peru (MINAGRI), which provided part of the data for this study and valuable information on how markets operates in the country, as well as to Rackham Graduate School and the Economics Department of the University of Michigan for their financial support in this project.

date of implementation, a fact which may suggest strategic behavior by market agents. In order to rationalize these empirical observations, we develop a Ramsey model of commodity taxation following Diamond (1975) which accounts for heterogeneity in household types and their distinct preferences over goods available in both market types (formal and informal). Finally, we will use the findings to recalculate VAT incidence taking into account the existence of informal sectors in the economy.

1 Introduction

Indirect taxes imposed on consumption, such as value-added taxes (VAT), are the primary source of government revenue in many developing countries (Jensen, 2022; Brockmeyer et al., 2024). According to a study by the Organization for Economic Cooperation and Development (OECD, 2022), taxes on goods and services constituted nearly half of the total tax revenue in 26 Latin American Countries (48.4%) in 2021, compared with a third (31.9%) in OECD countries. In contrast, direct taxes such as personal income tax and corporate income tax are less common in developing countries, where a large proportion of the population works in the informal sector and therefore avoids formal tax obligations. Indirect taxes, on the other hand, can be easier to administer and more difficult to evade, as they are embedded in the prices of goods and services and are therefore collected by businesses rather than individuals. This makes indirect taxes an attractive source of revenue for governments that lack the resources to effectively monitor and enforce compliance with tax laws.

The consensus view on the incidence of uniform consumption taxes is that they are regressive (Lustig, 2018; Warren, 2008) because they impose a larger tax burden on low-income households compared to high-income households. These arguments have led to governments applying a reduced (or zero) rate on necessity goods, such as food items. This is not optimal under the Atkinson and Stiglitz (1976) framework, according to which, under certain assumptions on utility function[1], it is better to avoid the distortions that arise from differential commodity taxation and redistribute using a non-linear income tax. However, recent efforts (Bachas, Gadenne, and Jensen, 2023) find that when informality is accounted for, uniform consumption taxes are progressive and reduce inequality in developing countries. The methods in this alternative literature rely on using the store type reported for each purchase as a proxy for household consumption from the informal sector, exploiting the differences across the household income distribution in the share of total expenditure in each store and relying on the important assumption that all modern (traditional) stores are formal (informal) with 100 (0) percent pass-through of taxes to consumer prices. However, there are numerous reasons to believe that the latter does not hold in reality. If this is the case, the question about the incidence of consumption taxes in developing countries remains unanswered.

[1]These assumptions are weak separability between consumption choices and income in utility, and homogeneity across individuals in the sub-utility of consumption. These two conditions imply that conditional on earnings, consumption choices provide no information on ability, and therefore differentiated commodity taxes create a tax distortion with no benefit.

One reason why the pass-through of taxes to prices is probably greater than zero in informal stores is related to nature of the VAT structure. Stores typically offset the tax they pay for their purchases with the tax they receive from their sales. Given that informal stores may be subject to paying VAT on some or all of their inputs (if they buy from formal stores), these taxes are not recovered through fiscal credit. Consequently, such taxes are regarded as an additional cost that must be transferred to consumers. This is part of the well known formal/informal dualism that makes incidence analysis intricate. A further rationale supporting the idea that pass-through is likely to be greater than zero relates to the possibility that traditional stores may charge equivalent or higher prices than their modern counterparts, even if informal. As a result, customers may end up paying more for identical products, instead of paying less, while the tax savings are kept almost entirely by the store. This phenomenon can be explained either by the informal stores confronting greater costs or by the existence of competitive forces with formal stores.

The assumption that the pass-through is 100 percent in formal stores does not seem to be accurate, either. When a final good is exempted from the VAT, taxes paid on inputs are in general not allowed to be used as fiscal credit (this is as opposed to a zero-rate tax, in which taxes paid on inputs can be recovered, or as opposed to a law that explicitly specifies that you can recover taxes paid on certain inputs of exempted goods[2]). Therefore, when a good is exempted, once again in this case taxes paid on inputs are regarded as costs, and we would expect the price to go down by less than the total amount of the tax, even in the case of perfect compliance. In fact, Benzarti, Garriga, and Tortarolo (2024) use high-frequency retail scanner data from Argentina and find that a temporary VAT cut had a 60 percent pass-through.

It is with this framework in mind that the current paper aims to address the question of what is the true incidence of the VAT in countries with substantial informality. This serves as a basis to explore the implications for optimal commodity taxation when the existence of informality is considered. We are going to address this question by looking at the case of Peru, which provides a convenient case study for numerous reasons. First, it has high levels of informality, with an informal employment rate of 71.6 percent in 2023[3], according to the International Labour Organization (2024), and with 58 percent of formal private firms having indicated that they faced competition from unregistered or informal firms in 2023, according to a report from the World Bank's Enterprise Surveys (World Bank, 2024). Secondly, it has high-frequency pub-

[2]The latter is the case for the law that regulates the policy we exploit in this study.

[3]Calculated as the share of informal employment in total employment for Youth and Adults (ages 15+) based on LFS - Permanent Employment Survey.

licly available data on wholesale and retail prices at the market-product-variety level for large markets that sell most of their transactions (if not all) in an informal way[4]. Thirdly, we recently got data for some formal supermarkets that allows us to calculate pass-through of the formal sector. Lastly, a policy announced in mid-April 2022 temporarily exempted certain products of the basic consumption basket from the value added tax (VAT), originally at 18 percent, starting on May 1st, 2022 through July 31st, 2022[5]. The goal of the policy was to limit the inflationary impact generated by the international economic crisis. The law exempted the last stage of production only, but stated that a fiscal credit was to be issued for the amount of VAT paid for the main (but not all) inputs (either domestically bought or imported) used in the production of the exempted goods. This presents us with a natural experiment to estimate the pass-through of taxes to prices for the products subject to a tax change as a result of the policy.

The incidence analysis relies on detailed expenditure microdata from the 2022 National Household Survey of Peru (ENAHO, for its spanish acronym). One of the key strengths of the survey is that it collects information on the various types of stores where households purchase products, such as supermarkets, specialty stores, and street vendors. By classifying these stores into formal, partially informal, and informal categories, we are able to gain valuable insights into household consumption patterns across both sectors. This categorization enables us to construct a proxy for household consumption from the informal sector, which, together with the estimation of the pass-through of taxes to prices as outlined in the preceding paragraph, will let us conduct a comprehensive analysis of the impact of consumption taxes on inequality, taking into account the prevalence of informality in the economy.

The findings from our pass-through analysis suggest that the pass-through of taxes to prices in informal stores differs considerably by type of products. While for some products we find no pass-through, for others such as eggs the pass-through is roughly 100 percent. At the same time, households in the bottom 10 percent of the income distribution spend, on average, 34 percentage points more of their income at informal markets than those in the top 10 percent of the income distribution.

One concern that might arise is how we decide which store is considered formal, partially informal, and informal, as well as how credible the assumption that the markets we have data on do not pay taxes on output sold is. In order to address

[4]That is, without remitting consumption taxes to the tax authority.

[5]There was a similar policy applied in 2023 that included more products to the exemption. We are in the process of getting updated data to test our findings with that new policy.

these concerns, we ran a survey for a subsample of these markets in Peru to determine what percentage of their sales are done in an informal way.[6] With the survey we also collect information about their price setting strategies which will serve as important inputs on the calibration for the production side of the model.

Our study adds to the body of literature that examines the impact of value-added taxes on the economy (Benzarti, Garriga, and Tortarolo, 2024), as well as that which acknowledges the influence of informality on the conventional conclusions drawn from previous research (Bachas, Gadenne, and Jensen, 2023). The primary contribution of this study is the estimation of pass-through for partially informal stores. Furthermore, this study distinguishes between pass-through analyses by product and formality status. The findings of this study challenge the commonly-held assumption in the literature (Bachas, Gadenne, and Jensen, 2023) that informal markets have a pass-through rate of zero, and accordingly, we recalibrate the incidence analysis by integrating our results.

The remainder of the paper is organized as follows. In Section 2, we discuss general aspects of the structure of the value-added tax and the institutional framework surrounding it in Peru. Section 3 outlines the details of the data used in this study. Section 4 and 5 elaborates on two alternative empirical approaches and present the results for each. Section 6 presents the setting for incidence analysis. Section 7 presents our model draft. Finally, in Section 8, we summarize our findings and outline the subsequent stages of this project.

2 Institutional setting

Value-added taxes (VATs) are taxes on goods or services levied at multiple stages of production, with the crucial feature of taxes on inputs being credited against taxes on output. They are ultimately borne by the final purchasers (International Monetary Fund, 2014)[7] and offer the advantage of securing revenue by the fact that they are collected throughout the process of production (unlike a retail sales tax) but without distorting production decisions (as a turnover tax does). In the absence of restrictions regarding crediting of taxes that are paid on goods and services used as inputs to productions (including those paid on investment goods), a VAT would be

[6] We collected information of 420 stands in 42 different markets in Lima, Callao, and the interior of the country.

[7] The real burden of the tax may not fall entirely on consumers but may in part be passed back to suppliers of factors through lower prices received by producers. We are going to set aside this possibility for the purpose of this paper.

economically equivalent to a pure retail sales tax. However, practical considerations, such as restrictions on crediting tax for investment goods and monitoring costs, make it challenging to maintain such equivalence.[8]

In Peru, our country of study, indirect taxes represented 60.1 percent of total tax revenue in 2021, with taxes on goods and services representing 48.4 percent of total revenue, 41.9 percent coming from general taxes and 6.5 percent from excise taxes (International Monetary Fund, 2023). Peru has a single rate of 18 percent for its VAT, with 2 percentage points corresponding to a tax towards the local authorities[9]. It is applied to all activities (with the exception of exports and certain products and operations that are exempted), such as buying and selling goods in the country, services, construction contracts, the first sale of real estate made by constructors, and imports of goods. It is collected through the invoice credit method, where each seller adds the applicable tax rate to each sale and issues an invoice indicating the tax charged. The buyer, if subject to VAT on their own sales, can offset the input tax paid on their purchases against the output tax charged on their sales. Any remaining tax balance is paid to the authorities, and excess credits can be claimed as a refund. The advantage of this method is that it is robust to the omission from taxation of any intermediate transaction. In cases where a vendor fails to levy or report a tax on a sale, any loss of tax revenue will be precisely compensated if the purchaser also omits to request the credit.

Exemptions in Peru are defined either in terms of particular commodities or particular regions. The rationale for exemption in general is related to ameliorating the distributional consequences of the tax, both through the effect on prices that consumers face, and through the effect on incomes, if we believe that producers of the exempted goods are on the lower end of the income distribution.

A considerable number of goods and services are exempt in Peru by the VAT law, including the sale or import of certain livestock, fresh vegetables, cereal, phosphates, fertilizers, etc. (see third column of Appendix Table B1 for a non-exhaustive list), as well as public transport services, cargo transportation, cultural performances, among others that are not relevant for the purpose of this study. These VAT exemptions do not give right of input VAT deduction, meaning that while the tax is not charged on output, tax paid on inputs cannot be reclaimed. Therefore, production decisions (and prices) may be affected by the VAT.

[8]See Brockmeyer et al. (2024) for a detailed description about how VAT works and its substantial importance in developing countries.

[9]Called "Impuesto de promocion municipal".

At the same time, the Amazon region has been granted a VAT exemption since 1999, with the enactment of Law 27037, as part of a government effort to promote development in the area. Despite the fact that this is not a zero-rating arrangement, as no credit is extended for taxes paid on inputs, it can be inferred that since the whole region is exempt, taxes on inputs are not levied. Consequently, this arrangement can be viewed as functionally similar to a zero-rating on output scheme, with consumption taxes exerting minimal influence on production decisions and prices.

The identifying policy variation we exploit is introduced by the Law 31452, which was issued on April 14th, 2022.[10] This law exempted specific food items that are part of the basic consumption basket from the VAT, both when coming from domestic or imported sources, with the aim of mitigating the inflationary impact generated by the ongoing global economic downturn. The exemption was in effect for three months, starting on May 1st, 2022 through July 31st, 2022, as depicted in Figure 1. The exempted goods comprised fresh, chilled, or frozen poultry meat (specifically, gallus domesticus), fresh eggs, granulated sugar, uncooked and unstuffed noodles, and bread. The reasoning behind including these goods was to alleviate the burden of rising food prices on low-income households, who typically spend a greater proportion of their income on these essential goods. At the same time, these are domestically produced goods that contribute significantly to the agricultural sector of the Peruvian economy.

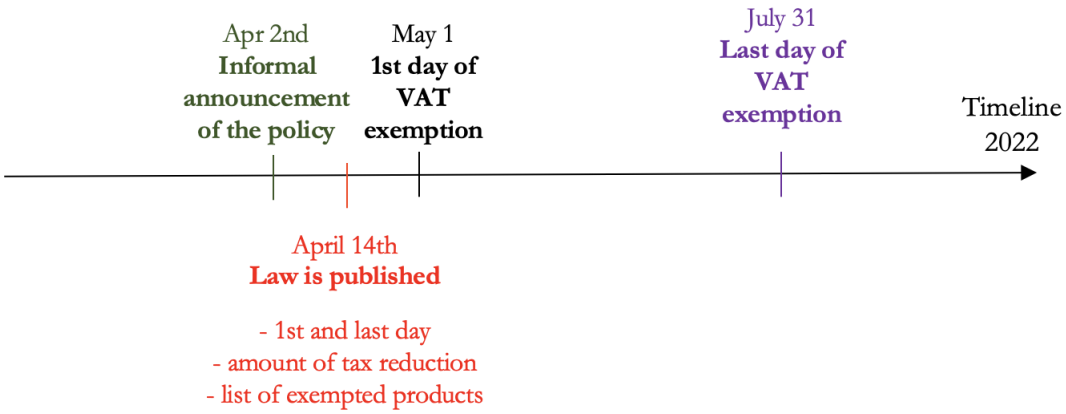
The policy under consideration established that a fiscal credit was to be issued for the amount of VAT paid for the primary inputs (but not all inputs) used in the production of the exempted goods, whether being bought from domestic sources or from abroad.[11] This arrangement constitutes a hybrid scenario that lies between exemption and zero-rating. Consequently, taxes will continue to affect production decisions and prices of these goods to a certain extent. In other words, pass-through could be anywhere between 0 and 100 percent for both formal and informal vendors.

The aforementioned policy offers a valuable opportunity to assess pass-through of taxes to prices in both the formal and informal sectors. Because of data constraints, a first version of this paper was focused on the informal sector, and followed the literature in assuming that the formal sector has a pass-through of 100 percent due to the salience of the tax, even though as stated earlier this is probably not accurate, even in the case of perfect compliance. However, we recently got data on high-

[10]The government announced on April 2nd that this law was going to be proposed, although there was no certainty then about the products to be included or the timeline of the policy.

[11]Columns 1 and 2 in Appendix Table B1 show the list of final products and inputs eligible for exemption for which we have data on prices.

Figure 1: Timeline of the policy



frequency prices data for a subset of products for some supermarkets (and we are waiting for additional data on a larger sample of products and stores), which would allow us to build the next version which will distinguish well the pass-through in the formal versus informal or partially informal sectors, as well as compare the results to those found in the existing literature (Benzarti, Garriga, and Tortarolo, 2024).

3 Data

3.1 Prices

Formal sector. We recently got access to high-frequency prices data for 80 varieties of products (including 3 out of the 5 exempted by the policy) sold by 76 supermarkets in Peru in 2022, the year of the policy, at the market level. In the first strategy we will explore, we do not really need this data since we can exploit differential VAT rates by products and regions, taking advantage of the regional exemptions mentioned in the previous section. Section 5, however, shows some preliminary evidence on the differential evolution of prices in formal markets versus partially formal markets during the policy period, and a future version will calculate pass-through accordingly.

Partially informal sector. The partially informal sector (hereby, informal sector, for simplicity) prices data comes from the "Domestic trade statistics of agricultural and agro-industrial food products" project, provided upon request by the Ministry of Agricultural Development and Irrigation of Peru, and elaborated by *Area de com-*

ercializacion de la Unidad Estadistica, *Oficina de Estudios Economicos y Estadisticos (OEEE)* and *Direcciones de Informacion Agraria Regionales y Subregionales*. It is collected with the goal of keeping agricultural producers and economic agents within the sector informed about the main aspects of the industry. It provides high-frequency data (depending on the geographic region and the market type, see below) for a total of 330 varieties of products at the market level, including 4 out of the 5 products that were provisionally exempted from the VAT under the policy[12]. It covers 61 wholesale markets and 67 retail markets across the nation, including the metropolitan area of Lima, 23 regional capital cities, and 3 subregional capital cities in Peru.[13] In order to favor comparability over time and across markets, people in charge of collecting the price information make sure that the product they choose meets the highest quality standards. They collect a sample of three to four prices for each product, and report the average in the data.[14] Daily data is available for wholesale markets in Lima, while the corresponding data for the interior of the country is available on Mondays, Wednesdays, and Fridays. Retail market data is available on Mondays, Wednesdays, and Fridays for both Lima and the interior of the country.

Some of these markets are publicly owned, and sellers entrance is governed by auctions, whereas others are privately owned and allot spaces for rent or sale to vendors. While wholesale markets are open to both commercial and retail customers, they are typically situated farther from urban centers, and consumers generally rely on nearby retail markets for their daily purchases. For the purpose of our analysis, however, we pool data from retail and wholesale markets. The reason we do this is that we want to maximize the number of observations in each regression, and because final consumers can actually buy from either.

Own data collection. The key premise underlying our analysis is that the majority of transactions in these markets are conducted outside the formal tax system. Even in the cases in which markets are publicly owned and have to comply with tax regulations, it is arguably safe to assume that there is a lot of space for informal transactions to occur, given the government's limited capacity for regular monitoring and enforcement. In order to test the accuracy of this assumption, we ran a survey

[12]These products are chicken, eggs, noodles, and sugar. None of the databases include price of bread, and the formal database also does not include noodles.

[13]The regional capital cities are Abancay, Arequipa, Ayacucho, Cajamarca, Cerro de Pasco, Chachapoyas, Chiclayo, Cusco, Huaraz, Huancavelica, Huancayo, Huanuco, Ica, Iquitos, Madre de Dios, Moqueagua, Piura, Pucallpa, Puno, Tacna, Tarapoto, Trujillo, and Tumbes. The subregional capital cities are Andahuaylas (in Apurimac), Jaen, and Chota (in Cajamarca).

[14]The stands within a market they collect the data from are mobile, meaning that they are not always the same each day.

in early May 2024 in 42 markets in our data in Peru (30 in Lima and Callao, 8 in Huánuco, and 4 in Ucayali) asking questions about types of products offered, price-setting strategies, sales volume, the level of competition both within the market and with formal stores, the extent to which clients pay in cash, credit card, or other electronic payment apps available[15], and other stand characteristics. Most importantly, we asked the surveyors to register if there were any clients arriving while the survey was conducted, as well as if they had been given a receipt. Additionally, we randomly selected stands in each market to purchase merchandise from (for a random amount between 5 and 10 soles, enough for these stands to be forced to give you a receipt according to the law) and registered whether a receipt was given. We ran 420 surveys (10 in each market), and made purchases in 311 of the surveys. A summary of results can be found on Appendix E. We find that out of the 318 purchases made, a receipt was given in only 5 of them (1.6%). Regarding the third party purchases that were made during the time the survey was being conducted, we found that in only 14 out of the 804 purchases a client was given a receipt (1.7%). We also observe that when given a receipt, the final price does not change. Therefore, we can safely assume that around 98.3% of the transactions in this markets are not reported to the tax authority, and that the price we observe in our data includes VAT.[16]

Appendix C shows anecdotal evidence and a scaled down version of what we found in the the survey. Two thoughts arise after our experience visiting these markets and looking at the receipts in Figures C4 and C5:

1. The final price did not change after asking for a receipt (it was the same as the price we see in the signs in the market). We can see that this includes VAT, so this is suggestive that the prices that we have in the data include VAT.

2. People usually do not get a receipt unless you ask for one. The receipt number in Figure C4 is 371, and the start date of these receipts that we can identify thanks to this tax payer id that we see (RUC - Registro unico de Contribuyente) is October 10th 2022. It is hard to believe that there were only 371 sales in almost 14 months. Similarly, the one on Figure C5 has 339 registered sales in almost 19 months. This is why we think that even though these stores are registered, a large part of their sales are informal.

Taking these two facts and the results in the survey together, we suspect that these

[15]The most popular apps of these kind in Peru are called Yape and Plin. These are comparable to Venmo in the US but with wider acceptance as a payment method in stores.

[16]The survey questionnaire is available upon request.

markets charge VAT in the price they post. Sometimes they report that VAT to the government, but most of the time they keep it to themselves.

National Household Expenditure Survey. We use data for consumption patterns across different store types from the 2022 National Household Expenditure Survey (in Spanish, "ENAHO"), which is conducted by the National Institute of Informatics and Statistics (in Spanish, "INEI") that has an expenditure module with a question that explicitly asks in which type of stores you buy each product, with information for 7550 food products in the second quarter of 2022, the module we use. We classify each purchase as "formal", "partially informal" or "informal" based on the store type each item is bought from, with traditional stores being classified as informal, the markets we have data on as partially informal, and modern stores as formal (see Appendix C for details on which store is classified in each group on Table C1, as well as a set of photographs that illustrate some types of stores/vendors). This gives us a proxy for food consumption from the informal sector at the household level, which we will later use in our next steps to construct measures of progressivity of the VAT. In particular, we look at the share of food expenditure across the income distribution, differentiating between the store types in which they are bought from, as well as whether the products are subject to the VAT cut or not.

4 Amazon vs. Non-Amazon Regions

4.1 Empirical strategy

In order to calculate the pass-through of taxes to prices in these markets, we adopt two different approaches. In this section, we explore the first approach, which involves comparing prices of the same policy-affected products in regions that are subject to VAT with those in regions that are never subject to VAT.

Thus, we restrict the analysis to products that are subject to the temporary VAT exemption. We classify observations in markets located in regions subject to tax (non-Amazon) as 'treated', and the same products in regions exempted from the tax (Amazon) as 'control'. This is because regions not subject to the tax should not respond to the policy. A clear advantage of this approach is that it allows us to ignore shocks that affect the international price of these products, or differences in seasonality by product, assuming these shocks affect both Amazon and non-Amazon regions in the same way.

'Treated' products refer to the final products that are exempt and for which we

have data: chicken, eggs, sugar, and noodles at informal stores, (see Table B1 in the Appendix for more details). The treated period we consider is determined by the law, spanning three months from May through July 2022.[17]

As a first step, we will provide graphical and non-parametric evidence of the pass-through by plotting the unconditional mean of the real price of products in the treatment and control groups. We look at the evolution of this measure before, during, and after the policy was implemented in May 2022. The assumption we rely on in order to calculate pass-through using this strategy is that goods in the treatment and control groups are reasonably similar and thus comparable. One way to test this assumption would be to see to what extent prices follow parallel trends in the absence of the law.

Next, we run the following regression that allows us to add standard errors and controls to the evidence provided by just looking at the unconditional means to precisely estimate the magnitude of the effect of the VAT cut on prices:

$$\ln p_{mpt} = \beta_0 + \beta_1 T_t + \beta_2 NonAmazon_{mp} + \beta_3 T_t * NonAmazon_{mp} + \alpha_m + \alpha_p + \varepsilon_{mpt} \tag{1}$$

where p_{mpt} is the price of product p in market m at date t, α_m and α_p are market and product FE, respectively, $NonAmazon_{mp}$ is equal to 1 if the product is at a market in the non-Amazon region (treatment group), and T_t is equal to 1 for the period in which the policy is applied. The coefficient of interest, β_3, estimates, on average, what is the percentage change in prices of treatment relative to control group means.[18]

Alternatively, we look at the corresponding coefficient for each product separately. This would be given by the coefficient β_3^p in the specification below, which is the

[17]Since we are dealing with informal stores that can arguably be motivated by speculation, we could also create alternative treated groups that include products that were initially thought to be included in the temporary VAT cut, as well as changing the treatment period to include those days/weeks in between the announcement and the implementation of the policy. We find that effects disappear after we include more products than the ones that are finally included in the policy or if we change the treatment period to include the window between announcement and implementation, which suggests that there was clear information about the products included and the timeline of the policy. We also look at results including inputs in the production of an exempted good (column 2 of Table B1). For some cases, we do see a reaction of the prices of these products, but this is not generally the case.

[18]Adding a time trend to control for each series having a trend overtime does not change results significantly.

same as the previous equation but excluding product FE and limiting the regression for the product we want to get the estimation for:

$$\ln p_{mpt} = \beta_0^p + \beta_1^p T_t + \beta_2^p NonAmazon_{mp} + \beta_3^p T_t * NonAmazon_{mp} + \alpha_m + \varepsilon_{mpt} \quad (2)$$

For each of these two regressions, we could two separate regressions: one that captures the effect of implementing the policy (from Jan2021 to Jul2022), and another one that captures the effect of ending the policy (from May2022 to Dec2022). Alternatively, we replace T_t by a categorical variable that takes the value of 0 for pre-policy period, 1 for policy period, and 2 for post-policy period. In the latter case, we would be comparing the pre-policy period with the policy period in one case, and the pre-policy period with the post-policy period in the other.

In all cases, we restrict the dataset to a panel of products that have fairly complete data in the period of analysis. As a way of dealing with outliers, we get rid of observations at the market-product-variety level that are farther than 5 standard deviations from the mean AND are higher than the 99th percentile of the standardized price. We also set as missing those observations that have a growth rate higher than 200 percent for one observation only and then go back to a similar level as before, since these are clearly data entry errors. After that we do linear interpolations to fill in missing values that appear either as a result of the previous outlier treatment or because there was no data collected at that date for that product-variety in that market. For most of our analysis, we use daily data, but similar results are obtained if we do the analysis by week instead, taking the weekly median per market-product-variety as our data point. We also pick one variety for each product based on data availability to avoid over-representation of products with too many varieties. Finally, we restrict the sample to those products that have at least one data point per week in the sample, and to those markets that have at least one observation per week for the period under consideration to make sure that the sample of markets we use for each product does not change each week.

4.2 Results

4.2.1 Pass-through analysis

Figure 2 shows the evolution of prices of the exempted products, capturing the mean across products of the median across markets by week, for treatment and control groups. In this specification, the treatment group is composed by those products that are temporarily exempted from the VAT with the law, in regions subject to the

Table 1: Regression results

	(1) Removal VAT	(2) Re-introduction VAT
(a) All products (pool)		
DiD coefficient (Non-Amazon vs Amazon)	0.0150	0.0799
	(0.0316)	(0.0513)
Observations	189,019	78,280
Pass-through rate		
(b) Eggs		
DiD coefficient (Non-Amazon vs Amazon)	-0.164**	0.0706***
	(0.0824)	(0.0176)
Observations	44,170	18,351
Pass-through rate	108%	

*Coefficients *100 show the change in price relative to t=0 (pre-policy period)*

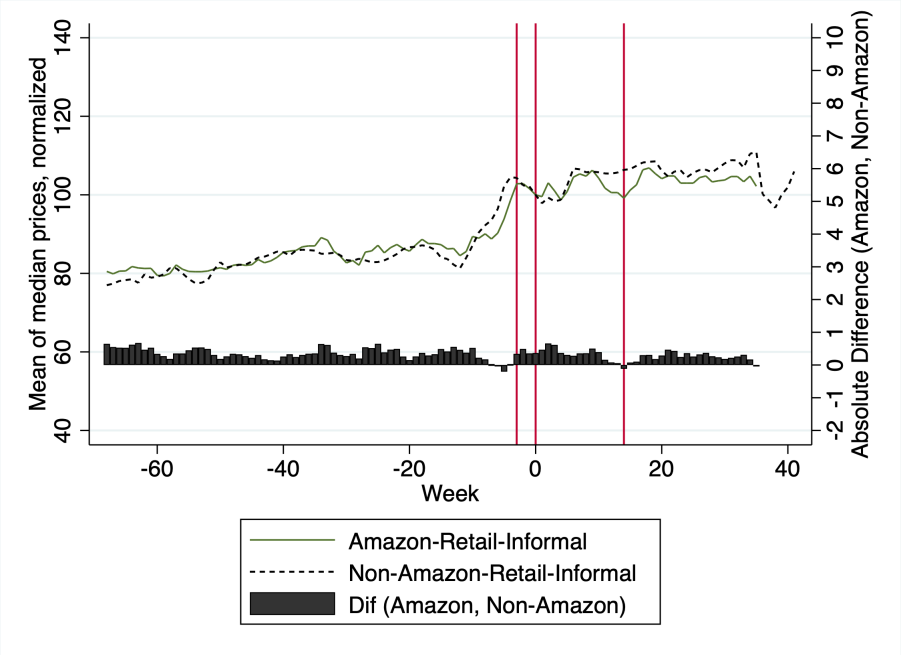
tax (that is, markets in non-Amazon regions). The control group is composed by the same products, but in regions that are not subject to the tax (markets in Amazon region). The first, second, and third vertical lines correspond to the day in which the law was announced, the first day of implementation, and the last day in which it had effect, respectively. Notice that this figure shows the values normalized so that for the week previous to the policy implementation the values are equal to 100. Since we do not see a large difference around the policy between the two groups for the pooled products, we look instead at the same picture by product.

On Figure 3 we see that the product in which there is a clear trend and a change affected by the policy is eggs. As Figure A1b shows, before the announcement of the policy, the treated group was consistently above the control group, in a relatively stable trend, in favor of the hypothesis that the assumption that goods in the control and treatment group are comparable. This relation is inverted during the policy window, where we see that the control group is above the treated group. After the policy is terminated, the pattern is less clear.[19]

Next, we look at the regression analysis using a Differences in Differences approach.

[19]Appendix A shows the results of these graphs for an alternative specification that corresponds to an event-study design similar to what we are going to see in the next section between formal and informal groups.

Figure 2: Evolution of prices for Amazon vs non-Amazon - pool of products



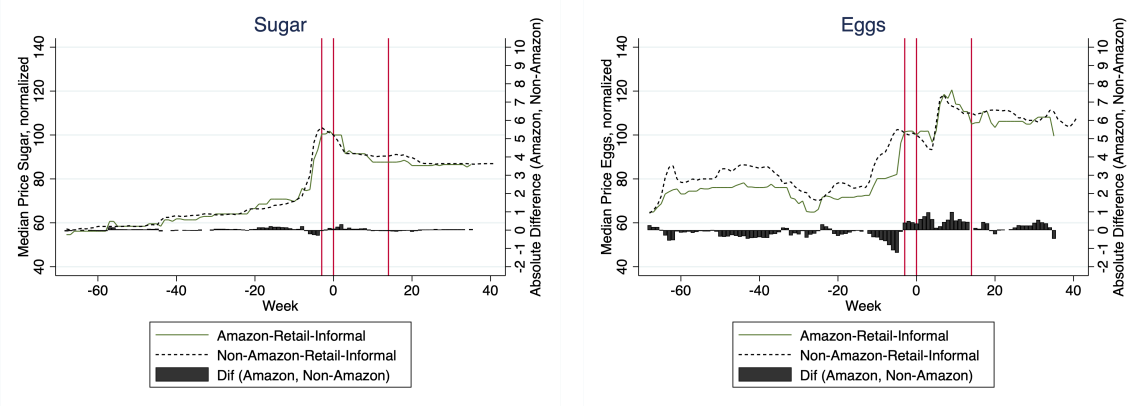
Source: own elaboration based on data from the Ministry of Agricultural Development and Irrigation of Peru.

Notes: the treatment group is composed by products subject to the policy change in regions subject to the tax (non-Amazon). The control group is composed by the same products in regions exempted from the tax (Amazon). Y-axis shows the median price of the pool of products, weighted by their importance in the consumption basket according to household surveys, normalized to 100 for the week before the policy is implemented (April 23rd-30th 2022=100). The vertical lines correspond to the dates of announcement of the policy, implementation, and final day, respectively.

Figure 3: Evolution of prices for Amazon vs non-Amazon - By product

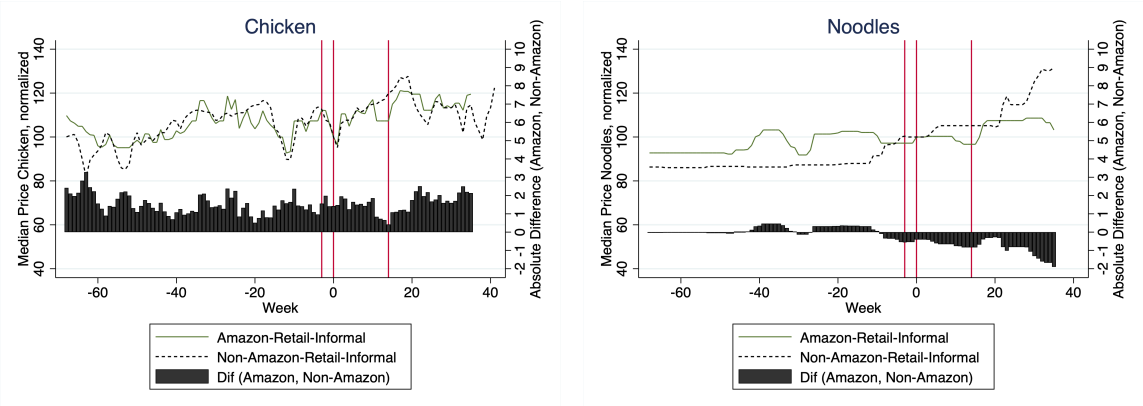
(a) Sugar

(b) Eggs



(c) Chicken

(d) Noodles



Source: own elaboration based on data from the Ministry of Agricultural Development and Irrigation of Peru.

Notes: the treatment group is the price of the exempted products in regions subject to the tax (non-Amazon). The control group is the price of the same product in regions exempted from the tax (Amazon). Y-axis on the left shows the median price of the product (across markets), normalized to 100 for the week before the policy is implemented (April 23rd-30th 2022=100). Y-axis on the right shows the absolute difference in price level (not normalized) between Amazon and non-Amazon. The vertical lines correspond to the dates of announcement of the policy, implementation, and final day, respectively.

Table 1 shows a summary of the results, reporting the correspondent coefficient β_3 for the pooled sample of products, as well as for the specification by product focusing on the case of eggs. As mentioned in the previous section, we split the analysis in two sub-periods, to capture any asymmetric effect between the implementation of the policy and its termination. Column 1 shows the coefficient looking at the implementation of the policy (that is, comparing pre-period with VAT exemption period) and Column 2 shows the coefficient for the reintroduction of the VAT (comparing pre-policy period with post-policy period).

We can see that if we take all four products, we find zero difference between prices in both groups. Our preferred specification is the one that looks at each product separately. We only show the case of the eggs because it is the product for which we could identify parallel trends before the policy was implemented. We see that there is a reduction of 16 percent in the price in treated regions relative to control regions, which for a VAT rate of 18 percent means the pass through for eggs was basically 100 percent. Looking at column 2 we can see that there is no difference between the price after the policy was removed and the pre-policy period, and we find that there is no statistical difference, consistent with the idea that the changes we see are caused by the policy. Tables D1 and D2 on Appendix D show all specifications.

5 Formal vs. Informal stores in Non-Amazon

5.1 Empirical Strategy

The second approach we use to calculate the pass-through of taxes to prices involves comparing products that are exempt from taxation under the policy with similar basic products that are not exempt, in regions that are always subject to VAT.

Thus, we restrict the analysis to regions that are always subject to the VAT (non-Amazon). We classify products subject to the temporary VAT exemption as 'treated', and a set of similar products that were not subject to the temporary VAT cut passed in April 2022—either because they were always exempted, zero-rated, or never exempted—as 'control'. A natural concern arises regarding how we define the correct set of products to use as the control group. We are currently working in this step, exploring products that are among the ten most consumed according to the 2022 ENAHO but are not exempt.[20] Additionally, an alternative strategy using synthetic

[20]The rationale behind this approach is that these products likely have similar elasticity to those included in the policy. We might take this one step further and analyze consumption patterns over several years, instead of only one year, to identify the right control group.

control groups or another machine learning method that helps match the evolution of prices for the pre-treatment period might constitute a better alternative; this is part of our next steps. We could also change the strategy to limit the set of non-treated products to those that are never exempted or those that were already exempted before the law was introduced, and therefore are not affected by the new policy.

Specific 'treated' products refer to the final products that are exempt and for which we have data on in both formal and informal markets: chicken, eggs, and sugar and the treated period spans three months from May through July 2022.

For this specification, we run the following regression to estimate the effect of the VAT cut on prices for treated and control products separately for both formal and partially informal stores:

$$\ln p_{mpt} = \gamma_0 + \gamma_1 T_t + \gamma_2 \text{treated}_{mp} + \gamma_3 (T_t \times \text{treated}_{mp}) + \alpha_m + \alpha_p + \varepsilon_{mpt} \qquad (3)$$

where p_{mpt} represents the price of product p in market m at date t, α_m and α_p are market and product fixed effects (FE) respectively, treated_{mp} equals 1 if the product is subject to the VAT exemption (treatment group), and T_t equals 1 for the period during which the policy is applied. The coefficient of interest, γ_3, estimates, on average, the percentage change in prices of the treatment group relative to the control group.

We expect the coefficient γ_3 to be larger in the regression for formal stores (F) than in the one for partially informal ($P.I.$) stores, i.e. $\gamma_3^F > \gamma_3^{P.I}$. If the stores classified as partially informal are fully informal, it could be the case that there is no impact of the VAT exemption on prices ($\gamma_3^{P.I} = 0$). If prices in partially informal stores do respond to the policy ($\gamma_3^{P.I} \neq 0$), this could indicate that these stores are more formal than originally thought (and thus, most of their prices reflect VAT compliance), or that they are adjusting their prices in response to competition with the formal markets, rather than directly responding to the VAT reduction.

In the survey we conducted to assess the extent of the informal operations in the partially informal markets, we found that in roughly 98% of the transactions a receipt was not given. Therefore, we would expect the influence of competition with formal stores to be more significant that the actual VAT reduction to explain their price setting.

5.2 Results

5.2.1 Pass-through Analysis

The Ministry of Agricultural Development and Irrigation of Peru already agreed on sending us a richer set of data for formal and informal markets that would allow us to construct a proper control group of similar products as the ones treated but that were not affected by the policy. Therefore, this version does not include the results of the exact exercise we want to do, but alternatively we wanted to see if there is any difference in the evolution of the prices of products in regions subject to tax, among formal versus informal stores, around the timeline of the policy. In order to do this, we run the following event-study design like regression:

$$\ln p_{mpt} = \lambda_0 + \sum_t \beta_t \text{week}_t + \sum_t \gamma_t \text{week}_t * \text{formal}_m + \alpha_m + \alpha_p + \varepsilon_{mpt} \tag{4}$$

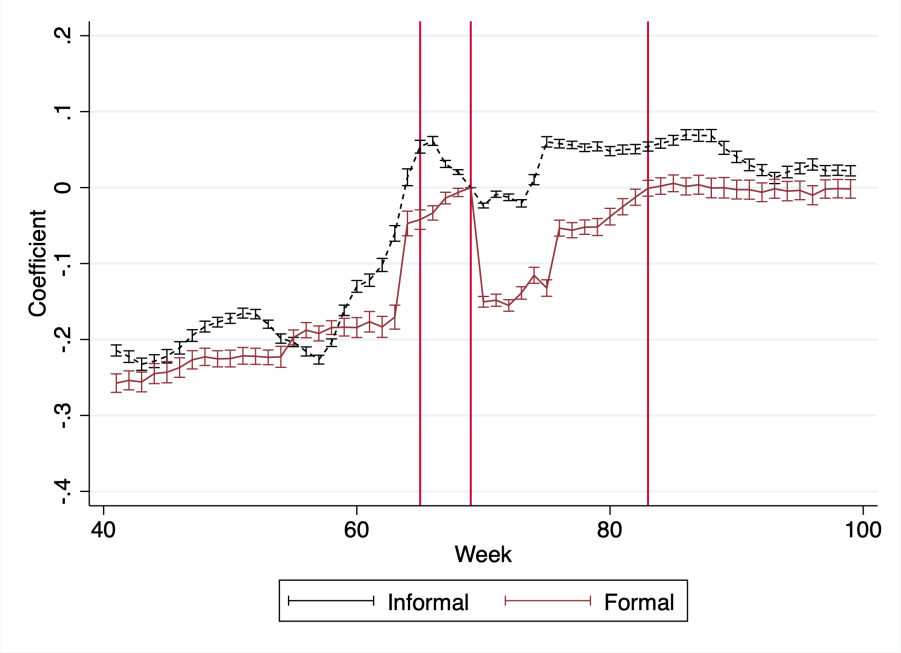
Where $week_t$ are dummies per week and the rest of the notation is similar to the previous equations. We restrict the analysis to 2022 because that is the period for which we have data on both formal and informal stores.

Each β_t reflects the mean across markets and products percentage change in price at week t for informal markets relative to the week right before the policy was implemented. $\beta_t + \gamma_t$ would be the equivalent for formal markets.

Figure 4 shows the β_t and $\beta_t + \gamma_t$ estimations with the corresponding confidence intervals, pooling all products together. We can see from this preliminary exercise that around the time the policy was implemented, the formal markets reduce their prices by roughly 16%, which would be consistent with the story of pass-through. The partially informal markets, however, begin a gradual reduction after the policy announcement, which continues with the actual implementation, but the total reduction is considerably lower than for the formal sector (less than 10%, as seen in the graph). The prices then continue to rise gradually, which could be the consequence either of less compliance with the policy or other factors that are raising the non-tax price level.

Figure 5 shows the same exercise by product. We can see that the same patterns repeat for each product separately: with the policy implementation, there is a sharp decrease in the price at formal markets. Informal markets, on the other hand, react earlier, after the policy announcement, in a gradual way and to a smaller extent than

Figure 4: Percentage change of prices at Informal vs. Formal markets, relative to the week before policy implementation



Source: own elaboration based on data from the Ministry of Agricultural Development and Irrigation of Peru.

Notes: these are the coefficients from equation (4) for a sample of products affected by the policy for which we have data in both formal and partially formal markets, in regions subject to the tax (non-Amazon). The vertical lines correspond to the dates of announcement of the law that regulates the policy, first day of implementation, and last day of the policy, respectively.

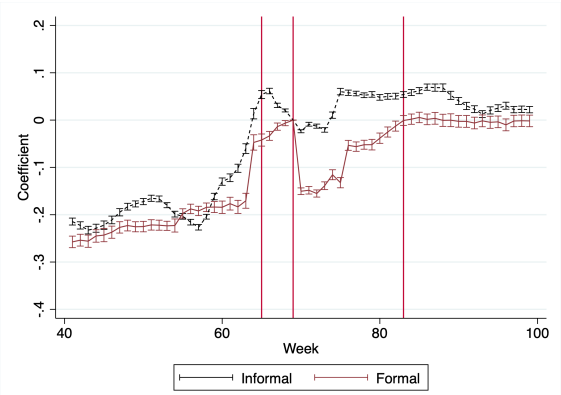
formal markets, but there is still a significant reduction in the price, consistent with the story that informal markets actually respond to a policy that changes the VAT even though most of their transactions are made in an informal way.

6 Incidence analysis

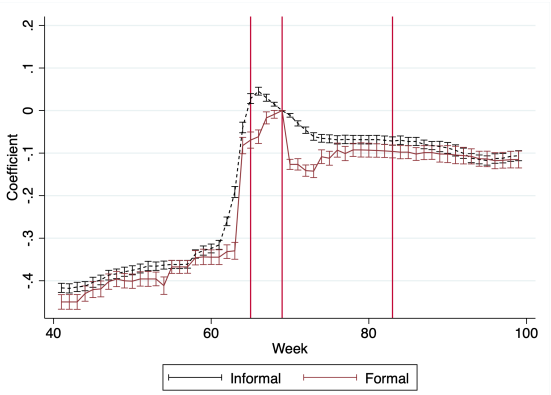
Regarding the distributional impact of the VAT cut, and considering that the policy was implemented to mitigate the negative impact of the international economic crisis on Peru's inflation, it is essential to examine household consumption patterns and estimate the share of food expenditure by store type across the income distribution. To achieve this, we use the National Household Survey (ENAHO) data

Figure 5: Percentage change of prices at Informal vs. Formal markets, relative to the week before policy implementation - By product

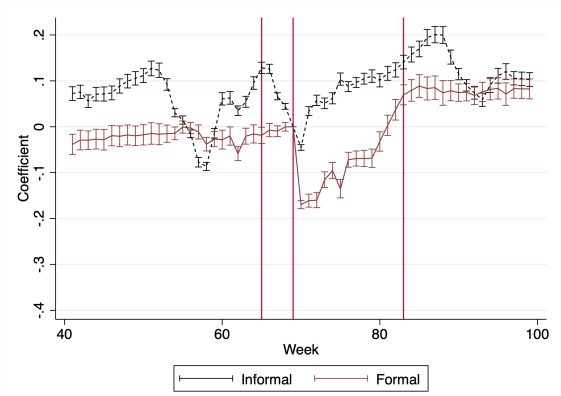
(a) All



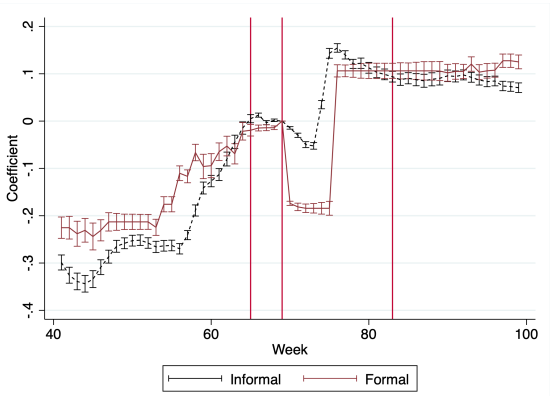
(b) Sugar



(c) Chicken



(d) Eggs



Source: own elaboration based on data from the Ministry of Agricultural Development and Irrigation of Peru.

Notes: these are the coefficients from equation (4) for a sample of products affected by the policy for which we have data in both formal and partially formal markets, in regions subject to the tax (non-Amazon). The vertical lines correspond to the dates of announcement of the law that regulates the policy, first day of implementation, and last day of the policy, respectively.

for the second trimester of 2022, a period that coincides with the announcement and first two months of implementation of the policy. In this survey, we calculate the food expenditure for both temporarily exempted and the rest of the food items (hereby "non-exempted"[21]), further breaking down the expenditure by different types of stores—supermarkets, partially informal markets, and fully informal stores—for each income decile[22].

Given the observed differences in pass-through rates between formal and partially formal (markets) stores, we can approximate the tax burden for each store category and, consequently, for households from different income levels. For formal stores, we anticipate a full pass-through, leading to a 15.2% decrease in final prices for VAT-exempt products. In contrast, the price reduction for partially informal stores depends on their response to price changes in formal stores. As illustrated in Figure 4, prices in informal stores decreased by approximately 9%. Thus, we assume the tax burden for consumers is 15.2% in formal stores and 9% in informal stores. For fully informal stores, we will let the pass-through be zero in one exercise, and 9% (as the partially informal stores) on another.

Figures 6 and 7 illustrate the share of exempted by the policy and non-exempted food items purchased in different types of stores by income decile. The exempted goods, affected by the VAT policy, include chicken, eggs, noodles, and sugar. The data shows that a significant portion of food expenditure occurs in informal stores (bottom bar, orange) and partially informal markets (medium bars, gray) across all income levels. This share ranges from 99% for the lowest income decile to 65% for the highest, decreasing as income rises, consistent with findings from previous studies (Bachas et al., 2023).

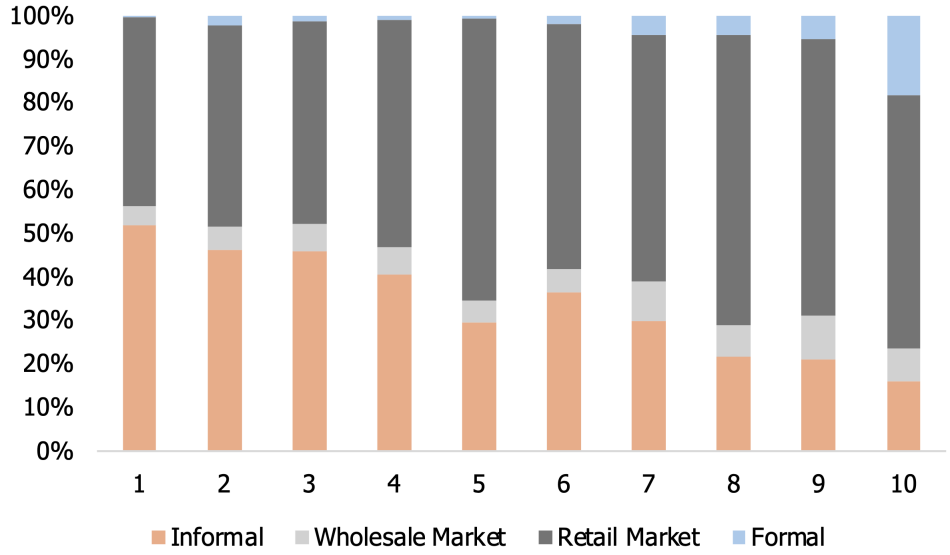
This pattern holds true for both exempted and non-exempted goods, though the share of consumption in markets is higher for exempted goods at all income levels. This is likely because exempted goods typically have low added value, resulting in similar quality across store types, making it more likely for individuals from all income levels to purchase these goods in the same locations (i.e., there is no significant benefit to shopping at a supermarket for these items).

Next we combine these shares and the pass-through results we found before to look at tax liability. We do a back of the envelope approximation of tax liability as the share of taxes that are paid by each decile, taking into account that the tax component in

[21] Although some of these goods include food items that are always exempted (mostly fruits and vegetables).

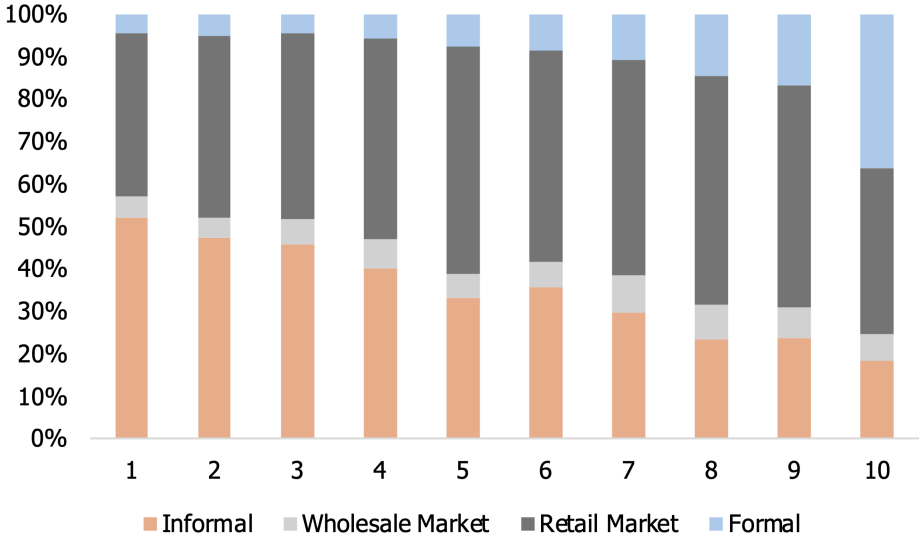
[22] See Table C1 on Appendix C for more details on store types and examples.

Figure 6: Share of Exempted Goods Expenditure by Store Type and Income Decile



Source: own elaboration based on data from ENAHO, 2022 April-June. Exempted goods include chicken, eggs, noodles and sugar.

Figure 7: Share of Non-Exempted Goods Expenditure by Store Type and Income Decile



Source: own elaboration based on data from ENAHO, 2022 April-June.

the expenditure we see will be different depending on how much each decile shops at each type of store. As we mentioned before, we assume full pass-through for formal stores, which for a 18% VAT rate corresponds to a 15.2% reduction in final price, and 9% reduction for partially informal stores, according to our earlier estimations. For fully informal stores, we present the results assuming zero pass-through on Figure 8a, and partial pass-through (9% reduction) in Figure 8b.

In Figure 8a we see that the tax liability for households in the lowest income decile is 4.3%, while it is 6.4% for the median household, and 8.7% for the highest decile. In other words, the VAT cut enables poorer households to retain 4.3% more of their income and wealthier households to retain 8.7% more. The blue line in Figure 8a illustrates the tax liability for exempted goods across all income deciles, suggesting that the VAT is more progressive than initially assumed once we take informality into account and under certain assumptions of the pass-through.

If we assume that the price response for non-exempted goods mirrors that of exempted goods (as shown in Figure 4), we can estimate the tax liability for non-exempted goods, represented by the dashed line in Figure 8a. The steeper slope of the tax liability curve suggests that the VAT on non-exempted goods is more progressive, meaning richer households bear a higher share of the tax burden. This aligns with the idea that non-exempted goods are less of a necessity and less standardized than exempted goods. Non-exempted goods are more likely to have added value, be sold at (formal) supermarkets, and be purchased by wealthier households.

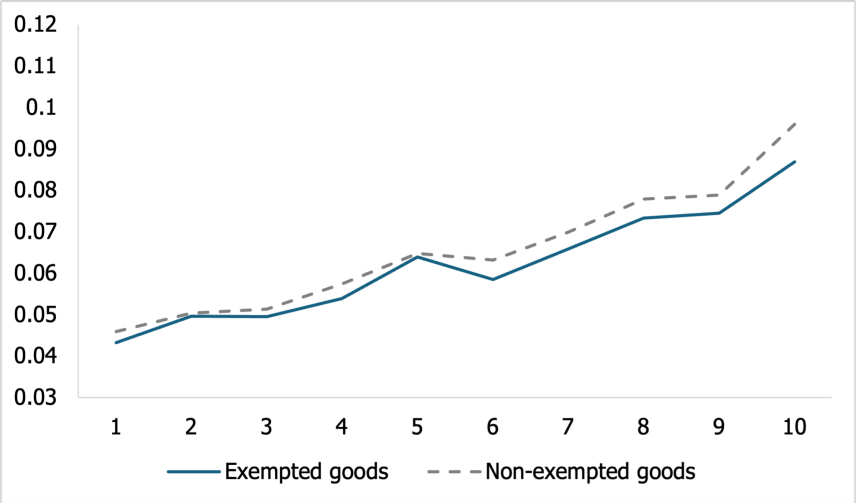
Interestingly, if instead of assuming zero pass-through for fully informal stores we assume the same pass-through as the partially informal stores that we found in our study, results change significantly, as shown in Figure 8b. The tax liability now looks flat, and only rises for the highest income decile, suggesting that even when we take into account informality, the fact that pass-trough is not zero at informal stores suggests that the burden of the VAT on household consumption is pretty similar in absolute terms across the income distribution. This is largely because informal stores raise their prices in response to price increases in formal stores due to VAT.

7 Model draft

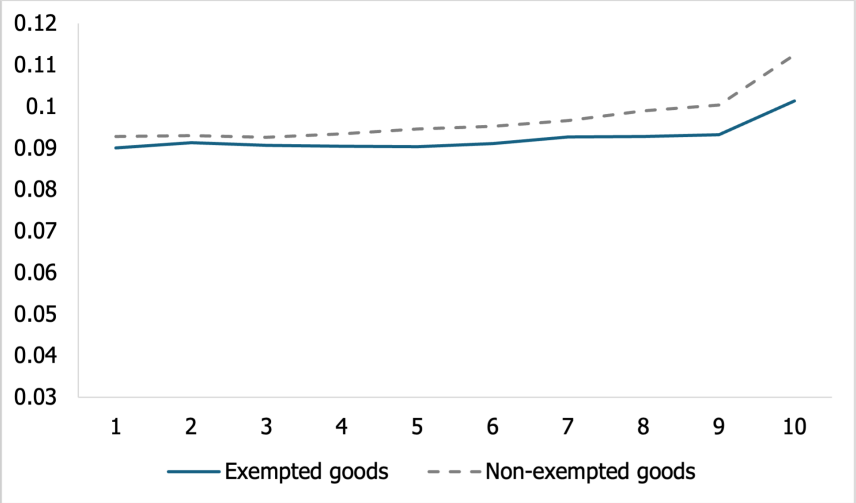
Having learned that pass-through differs across products, this section is devoted to developing a model that help us understand what determines the level of pass-through.

Figure 8: Tax Liability by Income Decile

(a) with zero pass-through at fully informal stores



(b) with partial pass-through at fully informal stores



Source: own elaboration based on data from ENAHO, 2022 April-June.
Notes: tax liability is approximated multiplying the expenditure at each store type by the corresponding pass-through. In both panels, we assume full pass-through for formal stores, which for a 18% VAT rate corresponds to a 15.2% reduction in final price, and 9% reduction for partially informal stores, according to our earlier estimations. For fully informal stores, we present the results assuming zero pass-through on Figure 8a, and partial pass-through (9% reduction) in Figure 8b.

We start with a Ramsey model of commodity taxation with heterogeneous households (Diamond, 1975), with a continuum of types that differ according to their level of income, in a context where there are j goods, and two varieties, formal and informal. The quality of the goods is equal in the formal and informal sector, with the difference that the amenities in each sector are different and consumption in the informal sector is not taxable. Households also differ in their valuation for the amenities.

Households. There are N households. Households (i) differ in their (exogenous) level of income (y^i). Each household has the same preferences over J goods, with each good available in both informal ($v = 0$) and formal ($v = 1$) varieties. The quality of the good will be the same in formal and informal varieties.

Both sectors offer amenities, though of different kinds. Formal stores provide conveniences such as accessible parking, cleaner facilities, air conditioning, and less crowded spaces—amenities generally preferred by wealthier customers. In contrast, informal stores offer flexibility with bulk splitting, allowing customers to buy smaller quantities—e.g., individual eggs, or single-use packets of detergent—instead of full-sized packages, and often provide an informal credit or tab system for trusted customers. These features make shopping more accessible and manageable for lower-income customers with limited cash flow.

Preferences over amenities while purchasing good j vary across households following an exogenous logit distribution A_j^i with pdf g and cdf G. Positive values of A_i mean that household prefer the amenities of the formal sector, negative values indicate a negative preference for the formal amenities, or equivalently a positive valuation of informal amenities. Zero indicates indifference.

To build intuition, we begin by assuming that the share of income household i allocates to good j, y_j, is exogenous. We assume that this share is the result from a prior optimization process in which each household determines how much income to allocate to good j. Thus, for each good, the only choice households face is the sector in which they will consume: either formal (c_{jF}) or informal (c_{jI}). We also assume that the good only consumes one good at a time, this will be good j.

For household i, if it purchases good j in the *formal* sector, the utility is:

$$U_F^i = u(c_{jF}^i) + A_i \tag{5}$$

where c_{jF}^i represents consumption of good j in the formal sector and $u(.)$ is twice

differentiable, increasing in consumption c^i_{jF} and concave. A_i denotes the valuation of formal-sector amenities. A positive A_i implies higher utility for households choosing the formal sector (i.e., the larger the A_i, the greater the utility from shopping formally).

Similarly, if the household buys good j in the *informal* market, the utility is:

$$U^i_I = u(c^i_{jI}) - A_i \tag{6}$$

where A_i now represents the valuation of informal-sector amenities. For those who prefer these amenities, A_i is negative, resulting in higher utility when shopping informally. $u(.)$ is twice differentiable, increasing in consumption c^i_{jI} and concave.

Therefore, household i's utility from consuming good j is as follows:

$$U^i = [u(c^i_{jF}) + A_i]v^i_j + [u(c^i_{jI}) - A_i](1 - v^i_j) \tag{7}$$

where v^i_j is an indicator variable representing the choice to buy good j in the formal sector.

Households face the following budget constraint:

$$p_{jI}c^i_{jI}(1 - v^i_j) + p_{jF}c^i_{jF}v^i_j \leq y^i_j \tag{8}$$

where p_{jI} (p_{jF}) is the consumer price of good j in the informal (formal) sector, and y^i_j is the exogenous income of household i allocated to the consumption of good j. The formal sector is taxable, while the informal sector is not. Let t_j be the tax rate on the formal variety of good j. Thus, consumer prices in both sectors follow:

$$p_{jF} = q_{jF}(1 + t_j)$$

$$p_{jI} = q_{jI}(1 + \mu t_j) \tag{9}$$

where q_{jF} (q_{jI}) is the producer price of good j in the formal (informal) sector. In the formal sector, the consumer price is the standard after-tax producer price. In the informal sector, the consumer price includes a mark-up associated with the tax rate. The term μ captures the indirect effect of tax rate changes in the formal sector on prices in the informal sector. If there is no indirect effect, then $\mu = 0$. If there is a positive strategic interaction in prices—meaning informal stores increase their prices when formal stores increase theirs due to a tax change—then $\mu > 0$.

For each good j, households choose whether to consume the formal variety ($v^i_{jF} = 1$)

or the informal variety ($v^i_{jI} = 0$) to maximize their utility given their exogenous amenities. The decision rule is thus:

$$v_{jF} = \begin{cases} 1 & \text{if} \quad u(c^i_{jF}) + A^i \geq u(c^i_{jI}) - A^i, \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

Given that the budget assigned to good j is already fixed, we can substitute the budget constraint directly into the utility function. Therefore, households will choose the formal variety of good j if:

$$\begin{aligned} U^i_F &\geq U^i_I \\ A_i &\geq \frac{1}{2}\left[u(c_{jI}) - u(c_{jF})\right] \\ A_i &\geq \frac{1}{2}\left[u\left(\frac{y_j}{q_{jI}(1+\mu t_j)}\right) - u\left(\frac{y_j}{q_{jF}(1+t_j)}\right)\right] \end{aligned} \tag{11}$$

u can be approximated as $u\left(\frac{y_j}{q} \cdot \frac{1}{1+\mu t_j}\right) \approx u\left(\frac{y_j}{q}\right) + u'\left(\frac{y_j}{q}\right)\left(\frac{y_j}{q} \cdot \frac{1}{1+\mu t_j} - \frac{y_j}{q}\right)$, which give us the final choice condition for buying good j on the formal sector:

$$A_i \geq \frac{1}{2} u'\left(\frac{y_j}{q}\right)\frac{y_j}{q}\left(\frac{1}{1+\mu t_j} - \frac{1}{1+t_j}\right) = A^* \tag{12}$$

where $\frac{1}{1+\mu t_j} - \frac{1}{1+t_j}$ represents the relative price difference between the informal and formal sectors due to taxes.

If $\mu < 1$, this term is positive, meaning that informal-sector prices are lower due to a reduced or avoided tax burden. For the household to choose the formal sector, A_i needs to be large enough to outweigh the utility benefits of lower prices in the informal sector. If $\mu > 1$, informal sector's prices are actually higher than those in the formal sector when taxes increase (i.e., an increase in the formal sector tax rate t_j leads to a greater-than-proportional increase in informal sector prices). The informal sector's price advantage diminishes or reverses. Only individuals with a stronger preference for informal amenities (more negative A_i) would continue to choose the informal sector despite the price increase.

As A_i is distributed following a logit distribution with cdf G, then the number of households that buy the good in the formal sector is given by $N_F = 1 - G(A^*)N$ and in the informal sector is given by $N_I = G(A^*)N$. This gives us the demand for good j in formal and informal sectors respectively.

Producers We assume a fixed number of firms, K , with a fixed share of formal and informal firms, represented by α_F and $\alpha_I = 1 - \alpha_F$. Each firm's productivity in producing good j is exogenous and given by ϕ_j^h. The production function is linear, $x_j^h = \phi_j^h$, implying that each firm has a set production level that cannot adjust to changes in demand or price.

Firms in the informal sector face an informality cost, R^h, representing the cost of getting caught, which varies by firm due to different potential implications if caught. This cost is exogenously given. Therefore, firm profits are defined as:

$$\begin{aligned} \pi_F^h &= \phi_j^h[q_j - c_j] \\ \pi_I^h &= \phi_j^h[q_j(1 + \mu_j t_j) - c_j] - R^h \end{aligned} \tag{13}$$

where q_j is the producer price of good j , c_j is the marginal cost of good j , R^h is the informality cost for firm h , and μ_j captures the strategic interaction in pricing between the formal and informal sectors, as previously defined.

Social Planner The social planner seeks to choose a tax rate t_j that maximizes total welfare, ensuring that the tax revenue generated from formal-sector consumption meets a spending target $\bar{g}$.

Since each household makes a binary choice to either consume formally or informally, welfare can be expressed as:

$$\max_t W = \int_{A^*}^{\infty} \left(u\left(\frac{y}{p_F(t)}\right) + A \right) dG(A) + \int_{-\infty}^{A^*} \left(u\left(\frac{y}{p_I(t)}\right) - A \right) dG(A) \tag{14}$$

where $p_F(t) = q_F(1 + t)$ is the formal sector price; $p_I(t) = q_I(1 + \mu t)$ is the informal sector price, which includes any indirect effect of the tax rate t_j through μ; A^* is the threshold value of A_i at which households are indifferent between the formal and informal sectors.

To meet the government's revenue target $\bar{g}$, the tax revenue from households in the formal sector must satisfy the following condition:

$$t\, p_F(t)\, y \cdot (1 - G(A^*)) \geq \bar{g} \tag{15}$$

Therefore, the optimal t_j will depend on the distribution of A, the utility function u, and the tax sensitivity in both the formal and informal sectors.

Market clearing conditions. Equilibrium requires that the total quantity supplied across both formal and informal varieties equals the total quantity demanded for good j. Given the fixed shares of formal and informal firms, represented by α_F and $\alpha_I = 1 - \alpha_F$, the market clearing conditions for each variety are as follows:

For the formal sector:

$$\sum_{i=1}^{N} v^i c_j^i = \alpha_F \sum_{h=1}^{K} x_j^h \tag{16}$$

For the informal sector:

$$\sum_{i=1}^{N} (1 - v^i) c_j^i = \alpha_I \sum_{h=1}^{K} x_j^h \tag{17}$$

Optimal VAT We introduce a Lagrange multiplier, λ, for the revenue constraint:

$$\mathcal{L} = W - \lambda \left[t, p_F, y \cdot (1 - G(A^*)) - \bar{g} \right] \tag{18}$$

From the first-order conditions, we obtain the following expression:

$$\begin{aligned} \frac{d\mathcal{L}}{dt_j} = &- \int_{A^*}^{\infty} \frac{y u' \left(\frac{y}{p_F} \right)}{p_F^2} q, dG(A) \\ &- \int_{-\infty}^{A^*} \frac{y u' \left(\frac{y}{p_I} \right)}{p_I^2} q\mu, dG(A) \\ &- \lambda \left[y \left(1 - G(A^*)\right) - t_j y g(A) \frac{dA^*}{dt_j} \right] = 0 \end{aligned} \tag{19}$$

If we assume $u(c) = \ln(c)$ and rearrange terms, the social welfare function becomes:

$$\begin{aligned} W = &[1 - G(A^*)] N [\ln(c_F) + \mathbb{E}[A \mid A \geq A^*]] \\ &+ G(A^*) N [\ln(c_I) + \mathbb{E}[-A \mid A \leq A^*]] \end{aligned} \tag{20}$$

The FOC is then:

$$\frac{1}{1 + t_j} - \frac{\mu\theta}{1 + \mu t_j} + \frac{\epsilon}{t_j} [\ln(c_F) - \ln(c_I)] = \lambda N y (1 + \epsilon) \tag{21}$$

where λ is the Lagrange multiplier (shadow price of government revenue); ϵ is the

elasticity of formal sector participation with respect to t_j; and $\theta = \frac{G(A^*)}{1-G(A^*)}$ represents the ratio of informal to formal sector households.

On the left-hand side (LHS), the first two terms represent the marginal welfare loss from a small increase in the tax rate t_j due to reduced consumption in the formal and informal sectors, respectively. The third term represents the marginal welfare effect due to households switching sectors. The right-hand side (RHS) represents the marginal value of government revenue from a small increase in t_j, adjusted for the change in the number of households in the formal sector due to sector switching.

We compare this FOC with the case where there is no strategic interaction between prices in the formal and informal sectors—equivalent to assuming households do not switch sectors when the tax rate changes. Here, the distribution of households between the formal and informal sectors remains constant, making μ irrelevant. In this scenario, total social welfare W simplifies to:

$$W = [1 - G(A^*)]N[u(c_F)] + G(A^*)N[u(c_I)] + \text{Constants} \tag{22}$$

Since A and $G(A^*)$ are constants in this case (no sector switching), we can treat the terms involving the informal sector as constants. Thus, W depends on t_j only through c_F. The FOC then follows:

$$\frac{1}{1 + t_j} = \lambda N y \tag{23}$$

In this FOC, compared to the one with strategic price interactions, we only have the marginal welfare loss from reduced consumption in the formal sector, as μ is assumed to be zero. The RHS represents the marginal value of government revenue from a small increase in t_j and does not include the $(1 + \epsilon)$ factor, as the tax base remains constant ($\epsilon = 0$).

This analysis indicates that without strategic interactions, the optimal tax rate (t_j^*) is higher compared to the scenario with strategic interactions. In the absence of strategic interactions, the tax rate only affects the formal sector consumption, as households do not switch to the informal sector regardless of tax changes. Consequently, there is no reduction in consumption within the informal market due to changes in prices in the informal sector (μ). Conversely, when strategic interactions are present, increasing the tax rate also induces welfare losses in the informal sector (in addition to those in the formal sector). These losses come from households switching to the informal sector, and the informal sector's prices adjusting upward

due to μ.

8 Next steps

We are enthusiastic about this project and look forward to improving the quality of the paper. We acknowledge that significant works remains to be done in order to address the research question, and we are actively working on it, as it remains a priority in our research agenda.

We are hoping the next version will include most of our future steps, including:

- Pass-through analysis using a differences-in-differences approach that compares the evolution for informal and formal stores for treated versus untreated products (we are waiting on a richer dataset to construct a proper control group)
- Finish the model, calibrate it, and look at the implications for optimal taxation of this partial pass-through in informal stores.
- Examine tax incidence using household surveys more deeply (explore Engel curves, Gini, etc.)
- Test if results hold when using similar policies implemented at different points in time for a different set of products

References

Atkinson, A.B. and J.E. Stiglitz (July 1976). "The design of tax structure: Direct versus indirect taxation". In: *Journal of Public Economics* 6.1, pp. 55–75. ISSN: 00472727. DOI: 10.1016/0047-2727(76)90041-4. URL: https://linkinghub.elsevier.com/retrieve/pii/0047272776900414 (visited on 11/30/2022).

Bachas, Pierre, Lucie Gadenne, and Anders Jensen (Sept. 21, 2023). "Informality, Consumption Taxes, and Redistribution". In: *The Review of Economic Studies*, rdad095. ISSN: 0034-6527. DOI: 10.1093/restud/rdad095. URL: https://doi.org/10.1093/restud/rdad095 (visited on 08/02/2024).

Benzarti, Youssef, Santiago Garriga, and Dario Tortarolo (Mar. 1, 2024). *Can Vat Cuts and Anti-Profiteering Measures Dampen the Effects of Food Price Inflation?* Rochester, NY. URL: https://papers.ssrn.com/abstract=4762948 (visited on 08/02/2024).

Brockmeyer, Anne et al. (Feb. 2024). "Does the Value-Added Tax Add Value? Lessons Using Administrative Data from a Diverse Set of Countries". In: *Journal of Economic Perspectives* 38.1, pp. 107–132. ISSN: 0895-3309. DOI: 10.1257/jep.38.1.107. URL: https://www.aeaweb.org/articles?id=10.1257/jep.38.1.107 (visited on 08/02/2024).

Diamond, P. A. (Nov. 1, 1975). "A many-person Ramsey tax rule". In: *Journal of Public Economics* 4.4, pp. 335–342. ISSN: 0047-2727. DOI: 10.1016/0047-2727(75)90009-2. URL: https://www.sciencedirect.com/science/article/pii/0047272775900092 (visited on 08/02/2024).

International Labour Organization (2024). *"ILO modelled estimates database" ILOSTAT*. URL: https://ilostat.ilo.org/data/.

International Monetary Fund, ed. (2014). *Government finance statistics manual 2014*. Washington D.C: International Monetary Fund. 446 pp. ISBN: 978-1-4983-4376-3.

– (2023). *Government Finance Statistics*. URL: https://data.imf.org/?sk=a0867067-d23c-4ebc-ad23-d3b015045405.

Jensen, Anders (Jan. 2022). "Employment Structure and the Rise of the Modern Tax System". In: *American Economic Review* 112.1, pp. 213–234. ISSN: 0002-8282. DOI: 10.1257/aer.20191528. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20191528 (visited on 08/02/2024).

Lustig, Nora, ed. (2018). *Commitment to equity handbook: estimating the impact of fiscal policy on inequality and poverty*. Washington, D.C: Brookings Institution Press. 1 p. ISBN: 978-0-8157-3221-1.

OECD, et al. (2022). *Revenue Statistics in Latin America and the Caribbean 2022*. URL: https://doi.org/10.1787/58a2dc35-en-es..

Warren, Neil (June 26, 2008). *A Review of Studies on the Distributional Impact of Consumption Taxes in OECD Countries*. Paris: OECD. DOI: 10.1787/241103736767. URL: https://www.oecd-ilibrary.org/social-issues-migration-health/a-review-of-studies-on-the-distributional-impact-of-consumption-taxes-in-oecd-countries_241103736767 (visited on 08/02/2024).

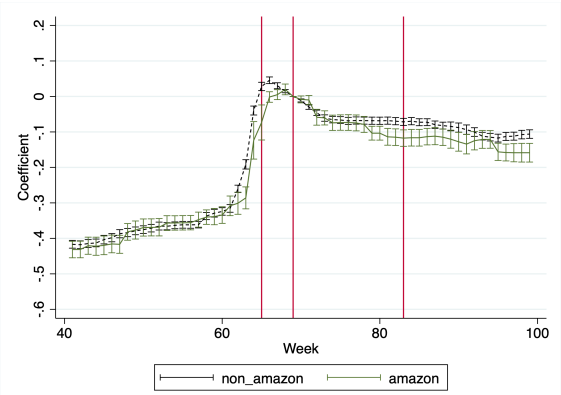
World Bank (2024). *Enterprise Surveys Country Profile: Peru 2023*. URL: https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/country/Peru-2023.pdf.

Appendices

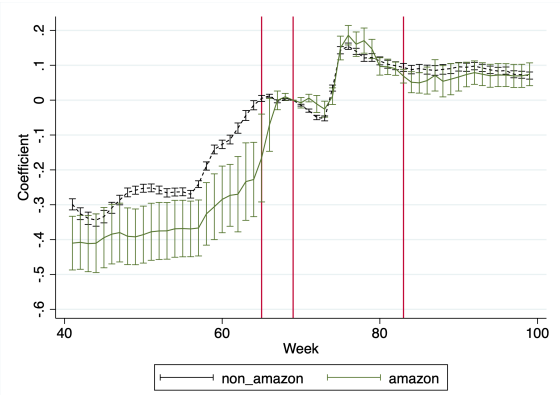
A Appendix A

Figure A1: Percentage change of prices at Amazon vs. Non-Amazon informal markets, relative to the week before policy implementation - By product

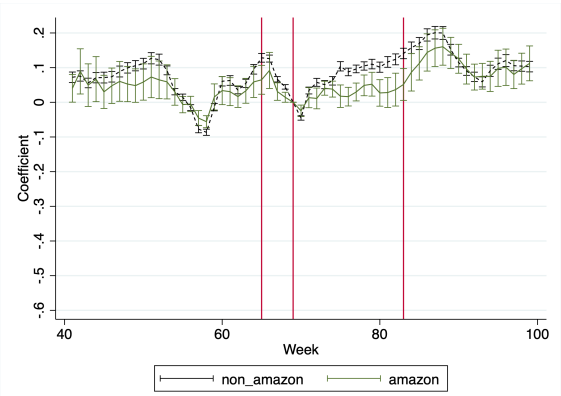
(a) Sugar



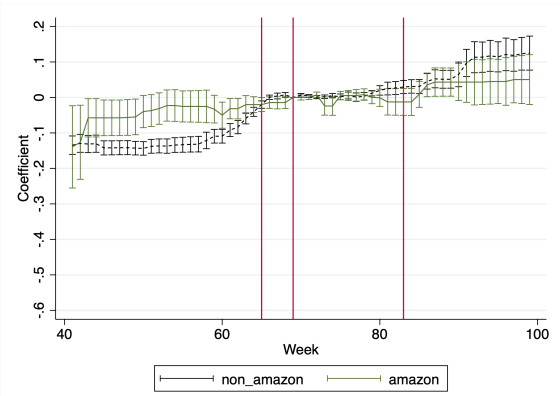
(b) Eggs



(c) Chicken



(d) Noodles



Source: own elaboration based on data from the Ministry of Agricultural Development and Irrigation of Peru.

Notes: these are the coefficients from equation (4) for a sample of products affected by the policy for which we have data in both amazon and non-amazon markets, in informal stores only. The vertical lines correspond to the dates of announcement of the law that regulates the policy, first day of implementation, and last day of the policy, respectively.

B Appendix B

Table B1: Product classification into treatment and control groups

Treated	Maybe Treated (Inputs)	Control		
Temporary VAT exemption	**Taxes on these inputs are allowed to be used as fiscal credit**	**Always VAT exempted**	**Standard 18% VAT (Never exempted)**	**VAT exemption in Amazon region**
Sugar (21 varieties)	Cooking oils (7 varieties)	Garlic	Quinoa	
Chicken	Flour (3 varieties)	Oats	Rice	
Pasta (5 varieties)	Chicken (poultry)	Onion	Beef	Same products as in Treated and Maybe Treated but in regions that are always exempted from the VAT
Eggs		Strawberries	Pork	
		Beans	Squash	
		Milk	Tuna	
		Lettuce	Corn	
		Lemon and limes	Turkey	
		Apples	Basil	
		Avocado	Bell pepper	
		...	...	
[28 product-varieties]	[11 product-varieties]	[176 product-varieties]	[99 product-varieties]	

C Appendix C

Table C1: Classification of store type reported in ENAHO as formal, informal, or other

Formal	Informal	Others
Bakery	"Ambulante" (street vendor/peddler)	Others
Supermarket	"Bodega" (traditional stores)	
Restaurants and bars	Wholesale and retail street markets	
	Foodtrucks/vans	
	Fairs	

Figure C1: "Informal" Markets in Lima, Peru

Figure C2: "Bodega"

Figure C3: "Ambulante"

Figure C4: "Start date: Oct-10-2022"

Comercial CORZO

COMERCIAL CORZO EIRL
MERCADO CENTRAL NRO. S/N OTR. CERCADO - TRUJILLO
RUC: 20481065012
FACTURA ELECTRÓNICA
F001-00000371

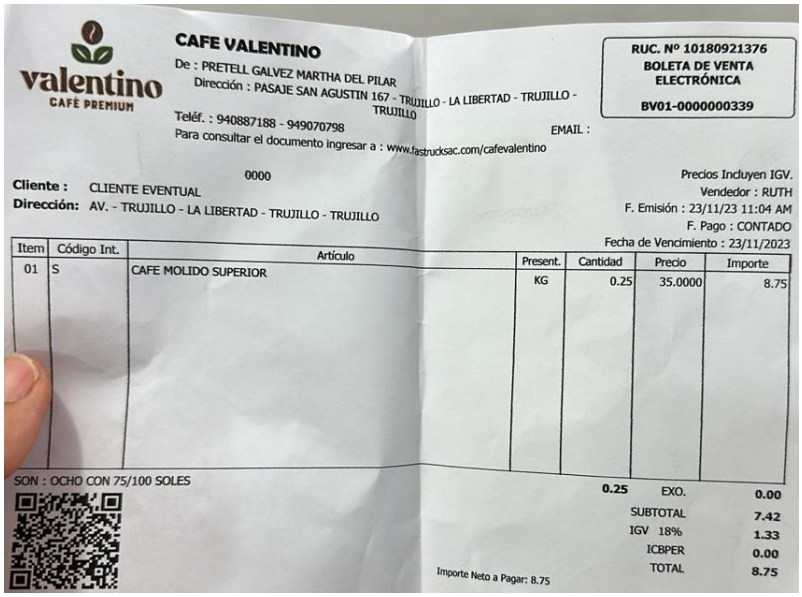
COD. VENTA: V0001494
FECHA: 23/11/2023 11:12 AM
R.U.C.: 20355267734
CLIENTE: DISTRIBUIDORA DESERET EIRL
DIRECCIÓN: JR. BOLOGNESI NRO. 666 LA LIBERTAD - TRUJILLO - TRUJILLO

DESCRIPCIÓN	CANT.	P.UNIT	IMPORTE
ACEITUNA EXTRA	0.50	20.00	10.00
QUESO MANTECOSO	0.25	32.00	8.00
MAIZ PERLA	0.25	12.00	3.00
LA SUELTA	0.05	100.00	5.00

OP. GRAVADAS:	S/	22.03
OP. EXONERADAS:	S/	0.00
OP. INAFECTAS:	S/	0.00
IGV 18%:	S/	3.97
ICBP (0):	S/	0.00
IMPORTE TOTAL:	S/	26.00

SON: VEINTISEIS SOLES CON 00/100
FORMA PAGO: Contado

Figure C5: "Start date: May-11-2022"



valentino CAFÉ PREMIUM

CAFE VALENTINO
De : PRETELL GALVEZ MARTHA DEL PILAR
Dirección : PASAJE SAN AGUSTIN 167 - TRUJILLO - LA LIBERTAD - TRUJILLO - TRUJILLO
Teléf. : 940887188 - 949070798
Para consultar el documento ingresar a : www.fastrucksac.com/cafevalentino
EMAIL :

RUC. Nº 10180921376
BOLETA DE VENTA ELECTRÓNICA
BV01-0000000339

0000
Cliente : CLIENTE EVENTUAL
Dirección: AV. - TRUJILLO - LA LIBERTAD - TRUJILLO - TRUJILLO

Precios Incluyen IGV.
Vendedor : RUTH
F. Emisión : 23/11/23 11:04 AM
F. Pago : CONTADO
Fecha de Vencimiento : 23/11/2023

Item	Código Int.	Artículo	Present.	Cantidad	Precio	Importe
01	S	CAFE MOLIDO SUPERIOR	KG	0.25	35.0000	8.75

SON : OCHO CON 75/100 SOLES

0.25	EXO.	0.00
	SUBTOTAL	7.42
	IGV 18%	1.33
	ICBPER	0.00
	TOTAL	8.75

Importe Neto a Pagar: 8.75

D Appendix D

Table D1: Regression results. Pre-policy vs policy

VARIABLES	(1)	(2)	(3)	(4)	(4)	(5)	(6)	(7)
	All four products				Sugar	Chicken	Eggs	Noodles
NonAmazon	0.0311	0.154***	0.0265	0.112***	0.112***	0.0553***	-0.0442***	-0.389***
	(0.0360)	(0.00509)	(0.0333)	(0.00873)	(0.00873)	(0.00403)	(0.0133)	(0.00765)
T	0.189***	0.189***	0.189***	0.263***	0.263***	0.0260	0.412***	0.0886**
	(0.0314)	(0.0314)	(0.0314)	(0.0541)	(0.0541)	(0.0247)	(0.0821)	(0.0420)
T*NonAmazon	0.0154	0.0137	0.0142	0.0611	0.0611	0.0162	-0.165**	0.0813
	(0.0316)	(0.0316)	(0.0316)	(0.0543)	(0.0543)	(0.0250)	(0.0824)	(0.0476)
Observations	189,019	189,019	189,019	69,797	69,797	63,624	44,170	11,428
R-squared	0.019	0.117	0.827	0.495	0.495	0.551	0.918	0.850
Market FE	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Product FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D2: Regression results. Pre-policy vs policy and pre-policy vs post-policy

Dependent variable: ln Price

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All four products				Sugar	Chicken	Eggs	Noodles
NonAmazon	0.0311	0.214***	0.0260	0.0712***	0.217***	0.0478***	-0.0262	-0.399***
	(0.0360)	(0.0141)	(0.0330)	(0.0220)	(0.0314)	(0.00764)	(0.0275)	(0.0156)
T_1	0.189***	0.189***	0.189***	0.189***	0.263***	0.0260	0.412***	0.0886**
	(0.0314)	(0.0314)	(0.0314)	(0.0314)	(0.0541)	(0.0247)	(0.0821)	(0.0420)
T_2	0.169***	0.169***	0.169***	0.169***	0.0895	0.127***	0.396***	0.128***
	(0.0511)	(0.0511)	(0.0511)	(0.0511)	(0.123)	(0.0266)	(0.0852)	(0.0328)
T_1*NonAmazo	0.0154	0.0147	0.0142	0.0150	0.0617	0.0176	-0.164**	0.0813
	(0.0316)	(0.0316)	(0.0316)	(0.0316)	(0.0543)	(0.0250)	(0.0824)	(0.0476)
T_2*NonAmazo	0.0794	0.0788	0.0791	0.0799	0.219*	0.0147	-0.0937	0.125**
	(0.0513)	(0.0513)	(0.0513)	(0.0513)	(0.123)	(0.0271)	(0.0854)	(0.0517)
Observations	236,409	236,409	236,409	236,409	87,401	79,654	55,251	14,103
R-squared	0.039	0.139	0.830	0.862	0.608	0.599	0.933	0.834
Market FE	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Product FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

E Appendix E

Table E1: Survey to Markets. Summary statistics

		Type of stand			
	All	Groceries	Animals and Eggs	Sausages & similar	Fruits and Vegetables
Share surveyed stands	100%	24%	32%	15%	29%
Average number of workers	1.60	1.51	1.81	1.53	1.46
Average number of customers in a day					
1-15	3%	1%	4%	0%	5%
16-30	27%	32%	26%	25%	24%
31-50	38%	33%	43%	39%	37%
51-100	22%	21%	16%	27%	26%
101+	11%	14%	11%	9%	7%
Average purchase value (in soles)					
s/1-5	3%	0%	2%	3%	5%
s/6-10	23%	17%	18%	20%	33%
s/11-20	42%	37%	34%	48%	51%
s/21-40	25%	39%	33%	17%	7%
s/41-60	4%	7%	4%	6%	1%
s/61+	4%	0%	8%	5%	2%
Mentioned "receipt" while describing a typical purchase	2%	4%	2%	3%	0%
Typical payment method					
Cash	64%	64%	64%	60%	66%
Credit Card	1%	0%	1%	1%	0%
Yape or Plin	36%	36%	35%	39%	34%
Made a purchase here	74%	84%	71%	80%	67%
% that were given a receipt (without asking for one)	1.6%	4.7%	0.0%	0.0%	1.2%
N other customers buying during survey	804	179	214	149	259
% other customers that were given a receipt	1.7%	2.2%	2.3%	0.0%	0.8%

Table E2: Survey to Markets. Price determination factors, by product

	Relevance of each factor on determining final price (0=nothing, 1=little, 2=relevant, 3=very relevant)																				Level of competition		
	Cost + Margin				Demand				Decide alone, looking at other stands in the market				Decide alone, looking at supermarkets				Decide together with other stands in the market						
Product	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	None	Little	High
apple	0%	28%	26%	47%	88%	9%	2%	0%	93%	7%	0%	0%	84%	12%	5%	0%	36%	36%	21%	7%	0%	30%	70%
avocado	2%	16%	41%	41%	91%	9%	0%	0%	93%	5%	0%	2%	95%	5%	0%	0%	30%	43%	20%	7%	0%	23%	77%
banana	0%	23%	38%	38%	83%	13%	4%	0%	94%	4%	2%	0%	85%	9%	7%	0%	32%	38%	17%	13%	0%	13%	87%
beef	0%	0%	100%	0%	100%	0%	0%	0%	100%	0%	0%	0%	100%	0%	0%	0%	20%	0%	80%	0%	13%	38%	50%
cheese	0%	21%	44%	35%	90%	7%	2%	2%	91%	5%	3%	0%	83%	16%	2%	0%	31%	43%	7%	19%	3%	21%	76%
chicken	1%	20%	39%	39%	82%	15%	2%	0%	92%	5%	2%	1%	88%	7%	5%	0%	32%	19%	14%	35%	4%	10%	87%
eggs	1%	18%	38%	43%	83%	9%	6%	2%	89%	10%	1%	0%	87%	13%	0%	0%	42%	22%	16%	20%	0%	12%	88%
fish	0%	0%	100%	0%	100%	0%	0%	0%	100%	0%	0%	0%	100%	0%	0%	0%	100%	0%	0%	0%	0%	17%	83%
honey	0%	0%	100%	0%	100%	0%	0%	0%	100%	0%	0%	0%	100%	0%	0%	0%	100%	0%	0%	0%	0%	100%	0%
legumes	0%	0%	0%	100%	100%	0%	0%	0%	100%	0%	0%	0%	100%	0%	0%	0%	0%	0%	50%	50%	0%	0%	100%
noodles	2%	28%	39%	31%	36%	30%	20%	14%	85%	12%	3%	0%	94%	6%	0%	0%	88%	8%	3%	1%	1%	15%	84%
olives	0%	11%	50%	39%	67%	28%	6%	0%	94%	6%	0%	0%	78%	17%	6%	0%	17%	28%	17%	39%	0%	22%	78%
onion	0%	24%	44%	32%	84%	12%	0%	4%	88%	12%	0%	0%	80%	20%	0%	0%	16%	44%	20%	20%	0%	20%	80%
pork	0%	0%	33%	67%	67%	33%	0%	0%	100%	0%	0%	0%	50%	33%	17%	0%	0%	100%	0%	0%	17%	33%	50%
potato	0%	19%	43%	38%	81%	10%	5%	5%	95%	5%	0%	0%	90%	10%	0%	0%	14%	38%	24%	24%	0%	10%	90%
rice	2%	21%	51%	26%	83%	11%	5%	1%	91%	6%	3%	0%	90%	9%	1%	0%	35%	36%	11%	17%	0%	17%	83%
sausage	2%	13%	44%	40%	77%	17%	2%	4%	88%	10%	0%	2%	83%	15%	2%	0%	38%	33%	15%	13%	4%	27%	69%
sugar	1%	47%	28%	24%	87%	3%	10%	0%	93%	0%	7%	0%	93%	1%	6%	0%	35%	12%	36%	17%	2%	22%	76%
tomato	0%	26%	43%	30%	70%	30%	0%	0%	87%	9%	4%	0%	70%	22%	9%	0%	13%	48%	30%	9%	0%	22%	78%
turkey	0%	0%	100%	0%	-	-	-	-	100%	0%	0%	0%	100%	0%	0%	0%	0%	0%	100%	0%	0%	100%	0%
average	**1%**	**24%**	**40%**	**35%**	**78%**	**13%**	**6%**	**3%**	**91%**	**6%**	**2%**	**0%**	**88%**	**10%**	**3%**	**0%**	**39%**	**28%**	**17%**	**16%**	**2%**	**19%**	**80%**

treated (2022 policy)