Documentos de trabajo Número II – 2022

Amazon's Effect on Prices: The Case of Mexico

Raymundo M. Campos Vázquez

Alejandro I. Castañeda Sabido

Aurora Ramírez Álvarez

Carlos Daniel Ruiz Pérez

El Colegio de México

agosto de 2022

http://cee.colmex.mx/documentos-de-trabajo

Amazon's Effect on Prices: The Case of Mexico[1]

Raymundo M. Campos Vázquez
Alejandro I. Castañeda Sabido
Aurora Ramírez Álvarez

El Colegio de México

Carlos Daniel Ruiz Pérez

July 2022

Abstract

Amazon's efficient logistics chain lowers costs, which may result in lower prices for consumers on and off its website. In this paper, we analyze whether Amazon's presence in Mexico has had a pro-competitive effect in lowering brick-and-mortar prices. We find that the entry of Amazon could have had caused an important pro-competitive effect by reducing brick-and-mortar retail prices; and this was found on the entry of Amazon platform but also to the product level: each time a product started selling through the Amazon platform, brick-and-mortar prices of that product decreased. The cities with the highest e-commerce penetration show the strongest decreasing price effects. We first study the Consumer Price Indexes (CPIs) for groups of products and cities. We find that cities with more e-commerce consumers tend to have lower CPIs in furniture and clothing, which are groups of products in which e-commerce has a greater presence. This difference is found to be important and statistically significant. We also analyze a novel database we created by merging average brick-and-mortar prices for a selected set of products (using public information from the National Institute of Statistics and Geography, INEGI) with information on products sold on Amazon (obtained through a third-party app, Keepa). We find that in general, when a product started selling on Amazon, brick-and-mortar prices decreased (though to different degrees, depending on the type of product). This pro-competitive decreasing price effect is observed with the introduction of products either sold and delivered by Amazon or sold and delivered by third parties on Amazon's website. Using a difference-in-differences empirical strategy, we find that this effect is more important in cities with larger numbers of e-commerce consumers. Finally, we conservatively estimate the welfare gains of this price decrease (ignoring other important welfare gains like increased product variety and increased ease of purchase) and find statistically significant welfare gains in states with higher absolute and proportional numbers of e-commerce consumers.

[1] This research was funded by project number 60831 of El Colegio de México. We appreciate the funding provided by Amazon for this project. Amazon did not edit any part of this document, suggest results, or share data. Replication files for academic purposes are available upon request.

Executive Summary

The aim of this paper is to analyze the competitive impact of Amazon's entry into Mexico. We gather information from private sources about the date on which Amazon started selling various products that have a major impact on the Mexican consumer price index (which includes 500 products). We match the characteristics of these products with Consumer Price Index information provided by INEGI. Then we measure the competitive effect of Amazon's sale of these products on the prices of brick-and-mortar stores, using a consumer index of prices based on individual prices from those stores.

- Our analysis confirms that e-commerce and brick-and-mortar retailers in Mexico operate in a single, highly competitive retail market. We find strong evidence that Amazon's entry has generated a significant pro-competitive effect by reducing brick-and-mortar retail prices and increasing product selection for Mexican consumers.
    - We find this effect to be significant when Amazon started selling these products in Mexico.
    - We also find this effect to be significant at the product level: every time a product was sold through the Amazon platform for the first time, brick-and-mortar prices of that same product decreased.
    - We find that prices decrease most in cities with the highest e-commerce penetration.
- We first examine the Consumer Price Index (CPI) by city and product group.
    - We find that, on average, cities with more e-commerce consumers tend to have lower CPIs in furniture and clothing, which are the product groups in which e-commerce may have more presence.
- We also analyze a novel database we constructed by merging the average prices of a selected set of brick-and-mortar products (obtained from public information from INEGI) with information on products sold on Amazon (obtained from a third-party app, Keepa).
    - We find that, in general, when a product was sold for the first time on Amazon, brick-and-mortar prices decreased, though in differing degrees depending on the type of product.
    - The effect of the entry of products sold and delivered by Amazon is a strong and statistically significant decrease in prices ranging from 1% to 28%.
    - The effect of the entry of products sold by third parties on Amazon is a decrease in prices ranging from 1% to 7% (except for an average 20% decrease in health products).
    - The most important effect is found on the first date on which the product was sold: there was a decrease in prices ranging from 1% to 24% for products sold by Amazon or by third-party sellers on the Amazon site.
    - These results are important because they show that the pro-competitive effect comes mainly from the product's listing on Amazon: it does not depend on whether the seller was Amazon or a third party selling on the Amazon site.
- A difference-in-differences (DID) analysis allows us to measure the effect of a new event. We define a treatment group (cities with more e-commerce consumers) and a control group, and we assume that the external event—the entry of Amazon—affects only the treatment group. We analyze the change in behavior between the groups, and we find that the pro-competitive effect is more important in the treatment group. The DID effects are significant both in the CPI aggregate analysis and in the product-by-product analysis.

- The DID effect in the CPI analysis of clothing and furniture is large. The CPI in cities with a high level of e-commerce users is 2.4 percentage points lower for clothing and 3.7 percentage points lower for furniture.
- The DID effect in the product-by-product analysis is a decrease ranging from 12% to 36%.
- We therefore conclude that the decreasing price effects of Amazon are much more important in cities with high levels of e-commerce users, which reinforces our hypothesis that the decrease in prices is caused by the entry of Amazon.

- Finally, we conservatively estimate the welfare gains of these price decreases (ignoring other important welfare gains like increased product variety and increased ease of purchase), and we find statistically significant welfare gains in states with greater numbers of e-commerce consumers.

Resumen Ejecutivo

El objetivo de este trabajo es analizar el impacto pro-competitivo que resulta de la entrada de Amazon a México. Obtenemos información de bases de datos privadas que nos dan la fecha en que Amazon empezó a vender productos que tienen un impacto significativo en el índice de precios al consumidor (500 productos). Compatibilizamos las características de estos productos con las de los productos incluidos en el índice de Precios al Consumidor publicado por INEGI. Posteriormente medimos el efecto pro-competitivo que genera la provisión de estos mismos productos por Amazon en los precios que fijan las tiendas físicas (el índice de Precios al Consumidor se obtiene de información de tiendas físicas).

- Nuestro análisis confirma que el comercio electrónico y las tiendas físicas compiten en un mercado altamente competitivo. El análisis encontró evidencia contundente de que la entrada de Amazon ha causado un importante efecto pro-competitivo al reducir los precios minoristas tradicionales.
    - Este efecto es estadísticamente significativo para la fecha de entrada de la plataforma de Amazon.
    - También es importante y significativo a nivel de producto: cada vez que un producto se lista en la plataforma de Amazon, los precios de ese mismo producto en establecimientos físicos caen.
    - Las ciudades con mayor penetración del comercio electrónico muestran reducción de precios más importante.
- Primero se analizan los IPC por grupo de productos y ciudades.
    - En promedio, las ciudades con más consumidores de comercio electrónico tienden a tener niveles más bajos de IPC de muebles y ropa, que son los grupos de productos en los que el comercio electrónico podría tener más presencia.
- Se analiza, además, una base de datos de precios promedio (de productos físicos) de un conjunto seleccionado de productos (de INEGI) con información sobre los mismos productos vendidos en Amazon (obtenida de Keepa).
    - En general, cuando un producto comienza a venderse en Amazon, los precios del mismo producto en tiendas físicas disminuyen (aunque en diferentes grados según el tipo de producto).
    - El efecto de la entrada de productos vendidos y entregados por Amazon es fuerte y estadísticamente significativo, con un rango de 1% a 28% de disminución en los precios.
    - El efecto de entrada de productos vendidos por terceros oscila entre 1% y 7% de disminución de precios (excepto el 20% de disminución promedio en productos de salud).
    - Sin embargo, el efecto más importante se encuentra en la primera entrada del producto (ignorando si el vendedor era Amazon o un tercero), que va desde un 1% hasta un 24% de efecto de disminución del precio.
    - Esto es importante porque muestra que el efecto pro-competitivo proviene principalmente del hecho de que el producto haya sido listado en Amazon, y no depende de si el vendedor fue Amazon o un tercero.
- Usando un modelo de DID (Diferencia en Diferencias es una técnica estadística que nos permite medir el impacto de la ocurrencia de un evento nuevo; para implementar la técnica se definen dos grupos de observaciones: un grupo de control y un grupo de tratamiento, y se asume que el evento externo —en este caso la entrada de Amazon— afecta solo al grupo de tratamiento y se analiza el cambio en comportamiento entre el grupo de control y el grupo de tratamiento), encontramos que este efecto pro-competitivo es más importante en las ciudades con más consumidores de comercio electrónico. El efecto DID es significativo tanto en el análisis agregado del IPC como en el análisis producto por producto.

- Los efectos DID en el análisis del IPC en ropa y muebles son importantes. El IPC promedio en ciudades con una alta proporción de usuarios de comercio electrónico es, en promedio, 2.4 puntos porcentuales más bajo en ropa y 3.7 puntos porcentuales en muebles.
- Los efectos DID en el análisis producto por producto oscilan entre 12% y 36%.
- Por lo tanto, se concluye que los efectos de Amazon en los precios de establecimientos físicos son mucho más importantes en ciudades con altos usuarios de comercio electrónico, lo que refuerza la hipótesis de que el efecto es probablemente causal: la caída de precios es causada por la entrada de Amazon.

- Finalmente, se estiman de manera conservadora las ganancias de bienestar de esta disminución de precios (es decir, se ignoran otras ganancias de bienestar importantes como la mayor variedad de productos, mayor facilidad de compra, etc.) y se encuentran ganancias de bienestar estadísticamente significativas en los estados con mayor cantidad de usuarios de comercio electrónico.

1. Introduction

The arrival of global retail e-commerce firms has caused a radical transformation in the ways that households spend their income. Not only have new entrants to retail e-commerce increased their presence, but traditional brick-and-mortar retailers have also done so online, and they too could be increasing the growth of e-commerce. New technologies have brought new ways to consume, which could be more efficient because they take advantage of important economies of scale and scope.

With the advance of technology in recent years, e-commerce has grown substantially in Mexico and the rest of the world. In Mexico, several important firms have taken positions in retail e-commerce, including Mercado Libre, Amazon, Walmart Online, Linio, and Alibaba.

Amazon has been able to create a successful logistics chain since it first entered Mexico. If we can show that Amazon exercises pricing over marginal cost (like almost any company), and that its prices have lowered those of their competitors, then the price evidence by itself would support the claim that it is more efficient. The aim of this study is thus to determine whether Amazon's entry into Mexico increased Mexican consumer welfare by offering lower prices than those of brick-and-mortar establishments.

We first study the Consumer Price Index (CPI)[2] by product group and city, and we compare cities with the highest and lowest shares of e-commerce consumers. We then conduct an event analysis using a novel database created by merging average prices of brick-and-mortar products obtained from INEGI with information on products sold by Amazon (obtained from a third-party app, Keepa, which keeps track of prices on Amazon Mexico). Our database allows us to analyze, on a product-by-product basis, the effect of the same or similar brick-and-mortar products appearing in the Amazon store.

Our results strongly suggest that the entry of Amazon into Mexico triggered an important pro-competitive effect and caused the lowering of brick-and-mortar prices. We find that cities with more e-commerce consumers tend to have lower levels of CPI per product group (note that the CPI in Mexico reflects brick-and-mortar products only). This difference is large and significant in furniture and clothing, as these categories include products that may have more presence in e-commerce.

Given that this CPI analysis could be too broad and that there could be other confounding factors (for example, the entry of other important e-commerce retailers), we further explore this hypothesis by using a microdata set to determine the effect of Amazon on a product-by-product basis. With these microdata we can control for various fixed effects. We find that, in general, when Amazon starts to sell a product, brick-and-mortar prices decrease (though to different degrees, depending on the type of product). The effect of products sold and delivered by Amazon is strong and statistically significant: a decrease in prices ranging from 1% to 28%. The effect of products sold by third parties on the Amazon website is not as strong and is not always statistically significant (especially when aggregated to the yearly level): a decrease in prices ranging from 1% to 7% (except for a 20% decrease in health products). Using a difference-in-differences (DID) empirical strategy, we also find that this decreasing effect on prices is more important in cities with larger numbers of e-commerce users.

---

[2] A Consumer Price Index is an economic indicator that measures, over time, the average variation in the price of a basket of goods and services representative of household consumption in Mexico. This means that it is a measure of prices from a specific bundle of products. It should be noted that in Mexico the CPI is based only on brick-and-mortar prices, except for electricity and telecom prices, which are the only prices obtained from websites; prices are traditionally obtained through on-site observation. The content of the basket is based on the National Household Expenditure Survey (ENGASTO) for 2012 and 2013 and the National Household Income and Expenditure Survey (ENIGH) for 2014. The weighting structure of the INPC generics considers household spending on consumer goods and services. In accordance with international recommendations, household spending in localities with less than 15,000 inhabitants is also included in the weighting structure. To achieve a greater spatial representativeness, prices were observed in nine additional cities or geographic areas, to reach a total of 55 geographic areas. This selection was made based on the economic, population, and political importance of the cities. For additional detail, see
https://www.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825104177.pdf

We quantify the effects of Amazon entry on local household welfare using these results as far as the data allow. We support our analysis on two stylized facts: i) Amazon entry occurs, theoretically, at the national level, as it does not require the presence of physical stores, but the deployment of a logistics chain (which increases in efficiency with the number of deliveries by taking advantage of economies of scale); and ii) the increase in Amazon users is gradual, and it increases unevenly in different cities and states, depending on their average income and degree of technological adoption.

These assumptions allow us to estimate the effect of Amazon's entry by comparing the cities with high and low levels of e-commerce consumers in 2020. The limitation of such an approach is that it underestimates the effect: cities with low levels of e-commerce could still have had a pro-competitive decreasing price effect caused by Amazon, to a lesser degree than those with high levels of e-commerce, but still positive. Our implicit assumption is that those with low levels of e-commerce were not affected by Amazon's entry, which results in the underestimation of effects in cities both with high and low levels of e-commerce. We thus obtain a conservative lower bound of the effects.

Finally, we make a rough conservative estimate of the welfare gains of this price decrease, ignoring other important welfare effects of e-commerce like increased product variety, increased ease of purchase, and elimination of commuting costs to brick-and-mortar establishments, and we find statistically significant welfare gains in states with large numbers of e-commerce consumers. Our estimations are very conservative and the resulting welfare effects are relatively small (though strongly significant). The effects are at least 0.06% of household income for clothing and 0.45% for furniture in 2020. They are, however, extremely important as they affect Mexico at the national level, and the pro-competitive effect will increase in time as e-commerce penetration is expected to increase exponentially in Mexico.

Our paper relates to other work that explores the economic consequences of the entry of foreign chains into developing countries (Iacovone et al. 2015; Javorcik & Li 2013). Most importantly, it is closely related to the literature that shows empirical evidence for these first-order effects of new entrant retail supermarkets in developing countries and the consequences of retail globalization (Atkin, Faber, & Gonzalez-Navarro 2018).[3] In contrast to previous studies, it focuses on the consequences for consumers in terms of the effect on prices, after the entry of Amazon brought about the growth of e-commerce in Mexico. Since Amazon is the major expanding e-commerce firm in our estimation period, our study also relates to the extensive literature on the effects of Amazon in the United States and other countries (He, Reimers, & Shiller 2021; Cavallo 2017; Goldfarb & Tucker 2019).

Cavallo (2017) conducted the first large-scale comparison of prices collected from websites and physical stores. His sample included 56 large multi-channel retailers in 10 countries, including Mexico.[4] He found that price levels were identical in about 72 percent of the comparisons; although price changes were not perfectly synchronized, they had similar frequencies and magnitudes. Online and offline price levels tend to have strong similarities and converge over time. If this is also true in Mexico, Amazon may also have a disciplining effect on the online prices of brick-and-mortar stores. However, we do not have information on those prices.

Our study helps to shed light on the effects on economy that is less developed but with growing e-commerce adoption, of the global technological innovation in retail. Our welfare analysis, rather than relying on cross-

[3] Atkin, Faber, and Gonzalez-Navarro (2018) obtained a new collection of Mexican microdata to estimate the effect of foreign supermarket entry on household welfare. They found that foreign retail entry (Walmart) causes large and significant welfare gains for the average household that are driven mainly by a reduction in the cost of living. About one-quarter of this price index effect is due to pro-competitive effects on the prices charged by domestic stores. They also show that the gains from retail FDI are on average positive for all income groups, but regressive. Finally, they found that the estimated gains are specific to foreign entry, rather than being driven by the entry of modern store formats more generally (with their event analysis controlling for fixed effects of municipality-by-barcode-by-month).

[4] The data he used was a match of online-offline datasets containing prices for more than 24,000 products, with 38,000 observations made from December 2014 to March 2016. This included some retail stores in Mexico, in particular Apple and Zara (clothing).

country trade flows (Fajgelbaum & Khandelwal 2014), price deflators (Topalova 2010), or simulated price changes (Porto 2006), uses price and consumption data at the household level and barcode-equivalent products to provide robust evidence of changes in the CPI. Thus, to the best of our knowledge, this is the first study to provide empirical evidence on first-order effects of e-commerce in Mexico.[5]

The rest of this paper is divided into five sections. In Section 2 we examine the CPI by city and product group, and compare the CPI of cities with higher and lower levels of e-commerce consumers. We also fit a DID regression to determine the statistical significance of this difference. In Section 3 we analyze a novel database we created by merging average prices from INEGI with price information on products sold on Amazon, and we analyze the price behavior in brick-and-mortar stores when a product starts selling on Amazon. We also use a DID regression to determine whether there is a more pronounced effect in cities with more e-commerce consumers. In Section 4 we estimate a very conservative approximation of the welfare gains to Mexican consumers attributable to the entry of Amazon, particularly in states with high levels of e-commerce. In Section 5 we summarize the results with some concluding remarks.

[5] Perhaps the only exception is Palacios (2001), who performs a summary of e-commerce in Mexico and qualitatively analyzes possible effects. However, 2001 was still too early to have found any effect.

2. CPI Trends: Cities with Low versus High Levels of E-Commerce

2.1. Data

Other things being equal, increasing the number of retailers that consumers can choose from is expected to increase competition and lower prices. An increased prevalence of e-commerce, which increases options for consumers, would therefore be expected to lower prices. This effect could be even more pronounced if the entrant is more efficient than existing retailers.

We seek to determine Amazon's effect on the prices of brick-and-mortar retailers in Mexico. Amazon is not the most important e-commerce seller (there are others, such as supermarkets), but we analyze Amazon because its data is available.

We first perform an exploratory analysis of the CPI, by city, of the main groups of products that Amazon sells online and for which the National Institute of Geography and Statistics (INEGI) has information.[6] INEGI collects prices to construct a nationwide Consumer Price Index, but its Mexican CPI does not include any information on e-commerce: all of its prices are obtained from brick-and-mortar establishments. This data is publicly available, and we gather information beginning in 2010. The information is available by product group (aggregated in a price index) or on a product-by-product basis (the observed price).

We also gather information from the ENDUTIH (National Survey on the Availability and Use of Household Information Technology 2020). The survey has information on the availability and use of information and communication technologies in Mexican homes and their use by individuals aged six years and older. We use state-level information regarding the number of people who buy goods on the internet (e-shoppers) during the last year.[7] We merge this information with the Consumer Price Index at the city level to obtain an estimate of the number of people who ordered online, which reflects the importance of e-commerce in the retail market.[8] It is important to note that INEGI includes only 55 large cities in the construction of the CPI.

On the one hand, the information from ENDUTIH helps us obtain a single snapshot, at the end of the analysis period (2020), of e-commerce users as a proportion of the total population of the state. As it is a survey, the lowest level of representative disaggregation is at the state level. On the other hand, the CPI reported by INEGI comes from prices gathered from the 55 largest cities in Mexico, so they are representative at the city level.

The ENDUTIH can help us to order the 32 states in Mexico from highest to lowest levels of e-commerce users. In each state, there are one to three cities for which CPI information is available. Though every city in a particular state will have the same e-commerce ranking, our analysis uses the city weights INEGI uses to aggregate cities to the national CPI, accounting for their relative importance. As the ENDUTIH is representative only at the state level, we use this information only to sort states (and, therefore, the cities within them) by their

---

[6] The INEGI data is organized by "Object of expenditure": "1. Food, beverages, and tobacco" (hereafter "food"); "2. Clothing, footwear, and accessories" (hereafter "clothing"); "4. Furniture, appliances, and household accessories" (hereafter "furniture"); "5. Health and personal care" (hereafter "health"); and "7. Education and recreation" (hereafter "leisure").

[7] For group 1 (food) we use the responses to "7.22 In the last 12 months, have you bought food and beverages online?" (P7_22_6). For group 2 (clothing) we use the responses to "7.22 In the last 12 months, have you purchased items for personal use, including clothing and accessories, online?" (P7_22_11). For group 4 (furniture) we use the responses to four questions: "7.22 In the last 12 months, have you bought household furnishings online?" (P7_22_12), "7.22 In the last 12 months, have you bought computers, laptops, or tablets online?" (P7_22_3), "7.22 In the last 12 months, have you bought cell phones or accessories online?" (P7_22_4), and "7.22 In the last 12 months, have you bought electronic devices online?" (P7_22_9). For group 5 (health) we use the responses to two questions: "7.22 In the last 12 months, have you bought personal hygiene, beauty, and cosmetic products online?" (P7_22_10) and "7.22 In the last 12 months, have you bought health items (medicines, braces, medical instruments, etc.) online?" (P7_22_16). For group 7 (leisure) we use the responses to three questions: "7.22 In the last 12 months, have you purchased photographic, telecommunications, or optical equipment online?" (P7_22_17), "7.22 In the last 12 months, have you bought books online (including e-books)?" (P7_22_1), and "7.22 In the last 12 months, have you bought video game consoles and games online?" (P7_22_15).

[8] Though the ENDUTIH has information on the people who ordered online (and even an approximation of their buying frequency), it does not have any information on prices or the amount spent. For this reason, we also analyze CPI data.

levels of e-commerce users (by type of goods bought). There are two approaches to defining the treatment and control groups (high versus low levels of e-commerce users) in the use of this data in the DID analysis. We can analyze states with a high versus a low number of total e-commerce users in absolute terms, or we can analyze states with a high versus a low proportion of e-commerce users, as a percentage of the total population.

Analyzing e-commerce penetration (that is, the proportion with respect to total population) could be more pertinent to analyzing the competitive effect of e-commerce: it could provide better insight into the percentage of the population able to compare brick-and-mortar to online prices, for which such competition would be more important. We would have the advantage of controlling by population: some small states could have a low number of buyers that are nonetheless a large part of their total population. However, we could also measure the importance of e-commerce in the retail market in absolute terms, as retailers could find the total number of buyers (even if it is a low proportion of the population) more relevant for business. For this reason, we present the results in absolute terms as a sensitivity analysis, in Appendix A. It should be noted, however, that the conclusions are qualitatively identical.

There are also various possibilities for the definition of "high" versus "low" levels of e-commerce:

i) We could choose a percentage of the population above which we would classify a state with a high level of e-commerce (e.g., more than 20% of population).
ii) We could compare only the five states with highest proportion of e-commerce consumers against the five with the lowest.
iii) We could classify the top half of the states as "high" and the bottom as "low."

The first approach would be ideal; however, we could find no guidance in the literature regarding what percentage to consider a "high" level of e-commerce users, especially in Latin America, where e-commerce penetration is not as high as in more developed countries. For this approach, we would have to arbitrarily define a percentage.

The second and third approaches are much simpler, as they merely divide the information at hand. We would not be deciding in an absolute sense what is "high or "low," only a relative use of e-commerce with respect to the rest of Mexico. However, there are two additional problems with the second approach: i) we would lose the data from states at the middle levels, and ii) the effects found could be much greater than the actual values, as we would be comparing only the tails of the distribution. For these reasons, we consider the third approach to be best.[9] We compare the 14 states with the highest proportion of internet consumers with the 18 states with the lowest proportion. This division allows us to keep a relatively balanced number of cities between the two categories, both by level of e-commerce and by product group.

There is another very important advantage to this approach. We are observing the proportion of e-commerce users only in 2020: a snapshot at that moment. We are therefore ignoring the possibility that the growth rate in e-commerce adoption over the period of our CPI analysis, from 2010 to 2020, could vary significantly from state to state. Our approach helps to minimize this problem: it maximizes the probability that states will remain in the same category of high or low levels of e-commerce throughout the period of analysis.

The disadvantage of this approach is that there might be states near the dividing line with almost the same degree of penetration, but in different categories. This can be seen in Table 3, where the percentages between the 14th and 15th places are usually very close or even the same. Our analysis, however, considers INEGI's city weights to account for the relative importance of each city. Those with the heaviest weights, which drive most of the results, are always Mexico City, Monterrey, and Guadalajara. These cities are in the group with a high level of e-commerce, as would be expected, no matter how we choose to divide the data.

[9] In 2018, INEGI included 55 cities in its CPI data. We use data from only 44 cities, as 11 of them were included only in 2015, so the time series is incomplete.

The validity of this procedure can be shown if we repeat the exercise but for 2015, the first year of the ENDUTIH. In 2015, as it was the first survey of its kind, there were no specific questions regarding the type of goods bought online;[10] however, there is a question about online purchases that we can use to compare the general growth in e-commerce from 2015 to 2020. The comparison in ranking of states is shown in Tables 1 and 2, in absolute and relative terms, for 2015 and 2020.

Table 1. Ranking of States by Number of Consumers Using Internet to Buy Goods, 2015-20

States	**2015**		**2020**	
	Users	**Rank**	**Users**	**Rank**
Aguascalientes	71,346	24	314,517	26
Baja California	329,614	6	980,848	6
Baja California Sur	56,267	29	246,011	27
Campeche	57,164	28	181,640	31
Coahuila de Zaragoza	143,829	14	698,316	12
Colima	52,101	30	224,683	30
Chiapas	89,887	20	371,956	24
Chihuahua	192,465	12	827,102	9
Ciudad de México	647,508	1	2,300,094	2
Durango	66,296	26	348,890	25
Guanajuato	223,769	9	935,768	7
Guerrero	84,205	21	376,638	23
Hidalgo	100,355	19	436,041	18
Jalisco	562,188	3	1,758,420	3
México	643,206	2	3,330,576	1
Michoacán de Ocampo	**132,181**	**15**	**609,255**	**14**
Morelos	75,371	23	384,047	22
Nayarit	65,124	27	243,141	28
Nuevo León	474,534	4	1,379,402	4
Oaxaca	75,491	22	396,999	20
Puebla	210,801	10	727,540	11
Querétaro	128,715	16	538,181	16
Quintana Roo	123,951	17	433,199	19
San Luis Potosí	110,040	18	484,812	17
Sinaloa	195,299	11	648,786	13
Sonora	259,842	7	734,702	10
Tabasco	70,533	25	393,760	21
Tamaulipas	258,445	8	871,897	8
Tlaxcala	30,060	32	179,837	32
Veracruz de Ignacio de la Llave	338,977	5	1,113,114	5
Yucatán	**158,455**	**13**	**558,858**	**15**
Zacatecas	48,371	31	227,362	29

Table 2. Ranking of States by Proportion of Consumers Using Internet to Buy Goods, 2015-20

States	**2015**		**2020**	
	Users	**Rank**	**Users**	**Rank**
Aguascalientes	6%	14	25%	11
Baja California	11%	1	29%	3
Baja California Sur	8%	7	30%	2
Campeche	**7%**	**13**	**20%**	**17**
Coahuila de Zaragoza	**5%**	**17**	**25%**	**12**
Colima	8%	8	31%	1
Chiapas	2%	32	7%	32
Chihuahua	**6%**	**16**	**23%**	**13**
Ciudad de México	8%	10	28%	4
Durango	4%	23	21%	16
Guanajuato	4%	20	17%	23
Guerrero	3%	29	11%	30
Hidalgo	4%	24	15%	24
Jalisco	8%	9	23%	14
México	4%	22	20%	20
Michoacán de Ocampo	3%	28	14%	28
Morelos	4%	21	20%	18
Nayarit	6%	15	20%	19
Nuevo León	10%	2	27%	5

[10] There is a similar question, but the groups of products are not equivalent to those used in the 2020 survey.

Table 2. Ranking of States by Proportion of Consumers Using Internet to Buy Goods, 2015-20

States	2015		2020	
	Users	**Rank**	**Users**	**Rank**
Oaxaca	2%	31	11%	31
Puebla	4%	25	12%	29
Querétaro	7%	12	27%	6
Quintana Roo	9%	4	26%	8
San Luis Potosí	4%	19	18%	21
Sinaloa	**7%**	**11**	**23%**	**15**
Sonora	10%	3	26%	9
Tabasco	3%	27	17%	22
Tamaulipas	8%	6	25%	10
Tlaxcala	3%	30	14%	27
Veracruz de Ignacio de la Llave	5%	18	14%	26
Yucatán	8%	5	27%	7
Zacatecas	3%	26	15%	25

As is clear from the tables, the growth rate of e-commerce varied among states, though all showed important increases, either in absolute or relative terms. In absolute terms, the growth rates are very similar, so they tend to maintain almost the same place from 2015 to 2020. Only two states, Michoacán and Yucatán, both on the dividing line between high and low, changed places (Michoacán would have ranked as a state with a low level of e-commerce in 2015 but with a high level in 2020). This means that only 6% of the states would be in a different group with our choice of division if the snapshot had been taken in 2015 instead of 2020. Only three cities (of 44) would be in a different group, also only a 6% difference. In relative terms, the growth heterogeneity is more evident, but only four states (12%) change group between 2015 and 2020 (Campeche, Coahuila, Chihuahua, and Sinaloa), and eight cities of 44 (18%).

If the division were wrongly defined, and states belonging in the group with a high level of e-commerce were mistakenly put in the group with a low level, or vice versa, the result would be less difference between the groups with low and high levels of e-commerce. Were the states correctly classified, the effect would be amplified, with greater differences, so wrongly defined states would only mean the results could be underestimated. In any case, 2020 is the only year of the ENDUTIH survey for which we can obtain information about the importance of e-commerce by product group.

In sum, the way we choose to divide states and cities with high versus low levels of e-commerce considers the fact that the growth rates of e-commerce could be different in some places than in others, but that most will remain in the same group even if we took the snapshot in a different year. This is much more important when we look at e-commerce users in absolute terms, so a sensitivity analysis using that approach is presented in Appendix A; it finds the same results.

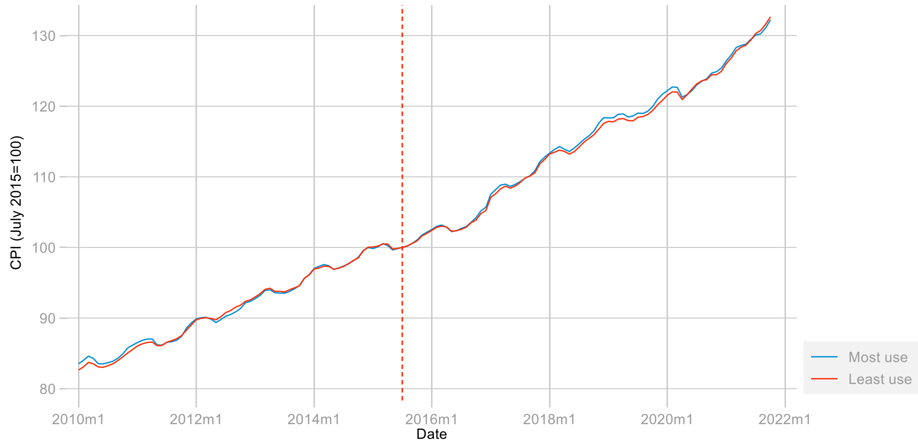
Table 3 shows the analysis in relative terms (cities with low levels of e-commerce users are highlighted in gray). The first column shows the percentage of e-commerce users with respect to the population of each state. Colima, Baja California Sur, and Baja California have the highest proportions: even if their population is not large in comparison to other states, a large proportion of their population consumes goods via e-commerce.[11]

[11] It is interesting that Colima, in relative terms, has the highest level of e-commerce, with 31% of its population making online purchases, while in absolute terms (shown in Appendix A) its rank is 30 of 32. The reverse is true of Estado de México, which ranks first in absolute terms but twentieth in relative terms, which places it in the group with low levels of e-commerce (except for food and leisure). Jalisco, which includes Guadalajara, the third-largest city in Mexico, ranks third or fourth in absolute terms, but fifteenth in leisure, which places it in the group with low levels of e-commerce for this category, though for other categories it remains in the group with high levels of e-commerce. In relative terms, Campeche is part of the group with high levels of e-commerce in the categories of food, health, and leisure.

Table 3. Cities by Ranking of States by Proportion of Consumers Using Internet to Buy Goods

City	General		Food		Clothing		Furniture		Health		Leisure	
	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank
Colima, Col.	31%	1	9%	7	22%	1	18%	2	13%	1	8%	5
La Paz, B.C.S.	30%	2	8%	10	19%	3	18%	4	11%	5	7%	7
Tijuana, B.C.	29%	3	9%	6	19%	2	19%	1	10%	6	8%	4
Mexicali, B.C.	29%	3	9%	6	19%	2	19%	1	10%	6	8%	4
Mexico City	28%	4	13%	2	17%	7	16%	6	11%	4	8%	3
Monterrey, N.L.	27%	5	11%	3	18%	4	18%	3	13%	2	9%	2
Querétaro, Qro.	27%	6	10%	4	18%	5	17%	5	10%	9	9%	1
Mérida, Yuc.	27%	7	16%	1	15%	15	16%	8	11%	3	7%	8
Chetumal, Q.R.	26%	8	9%	5	15%	14	16%	9	9%	11	8%	6
Hermosillo, Son.	26%	9	8%	9	16%	12	16%	7	9%	10	7%	9
Huatabampo, Son.	26%	9	8%	9	16%	12	16%	7	9%	10	7%	9
Matamoros, Tamps.	25%	10	5%	20	17%	9	14%	13	8%	16	4%	21
Tampico, Tamps.	25%	10	5%	20	17%	9	14%	13	8%	16	4%	21
Monclova, Coah.	25%	12	7%	12	17%	6	15%	10	10%	7	6%	11
Cd. Acuña, Coah.	25%	12	7%	12	17%	6	15%	10	10%	7	6%	11
Torreón, Coah.	25%	12	7%	12	17%	6	15%	10	10%	7	6%	11
Cd. Jiménez, Chih.	23%	13	6%	16	17%	8	14%	12	7%	21	6%	12
Chihuahua, Chih.	23%	13	6%	16	17%	8	14%	12	7%	21	6%	12
Cd. Juárez, Chih.	23%	13	6%	16	17%	8	14%	12	7%	21	6%	12
Tepatitlán, Jal.	23%	14	7%	11	16%	13	13%	14	9%	12	5%	15
Guadalajara, Jal.	23%	14	7%	11	16%	13	13%	14	9%	12	5%	15
Culiacán, Sin.	23%	15	5%	22	17%	11	13%	15	8%	14	5%	16
Durango, Dgo.	21%	16	5%	18	14%	17	12%	16	7%	17	4%	18
Campeche, Camp.	20%	17	8%	8	13%	19	12%	18	9%	13	6%	14
Cuernavaca, Mor.	20%	18	5%	19	14%	18	11%	21	7%	18	4%	20
Tepic, Nay.	20%	19	5%	21	14%	16	11%	19	7%	19	3%	26
Toluca, Edo. de Méx.	20%	20	6%	14	12%	21	12%	17	7%	20	6%	13
Villahermosa, Tab.	17%	22	5%	17	11%	23	11%	22	8%	15	4%	22
Cortazar, Gto.	17%	23	4%	24	11%	22	9%	26	6%	22	4%	17
León, Gto.	17%	23	4%	24	11%	22	9%	26	6%	22	4%	17
Tulancingo, Hgo.	15%	24	2%	31	10%	25	7%	29	4%	29	4%	24
Fresnillo, Zac.	15%	25	3%	27	11%	24	10%	23	6%	24	4%	23
San Andrés Tuxtla, Ver.	14%	26	4%	25	9%	28	9%	25	5%	25	3%	28
Córdoba, Ver.	14%	26	4%	25	9%	28	9%	25	5%	25	3%	28
Veracruz, Ver.	14%	26	4%	25	9%	28	9%	25	5%	25	3%	28
Tlaxcala, Tlax.	14%	27	2%	29	10%	26	9%	24	4%	28	3%	27
Jacona, Mich.	14%	28	3%	26	9%	27	8%	27	5%	26	3%	29
Morelia, Mich.	14%	28	3%	26	9%	27	8%	27	5%	26	3%	29
Acapulco, Gro.	11%	30	2%	30	8%	30	6%	31	4%	30	2%	31
Iguala, Gro.	11%	30	2%	30	8%	30	6%	31	4%	30	2%	31
Tehuantepec, Oax.	11%	31	3%	28	6%	31	6%	30	4%	31	2%	30
Oaxaca, Oax.	11%	31	3%	28	6%	31	6%	30	4%	31	2%	30
Tapachula, Chis.	7%	32	2%	32	5%	32	4%	32	3%	32	2%	32
Number of cities with high levels of e-commerce		21		18		22		21		19		19
Number of cities with low levels of e-commerce		23		26		22		23		25		25

Figure 1. Weighted Average of General City CPI, by E-commerce Customers per Capita.

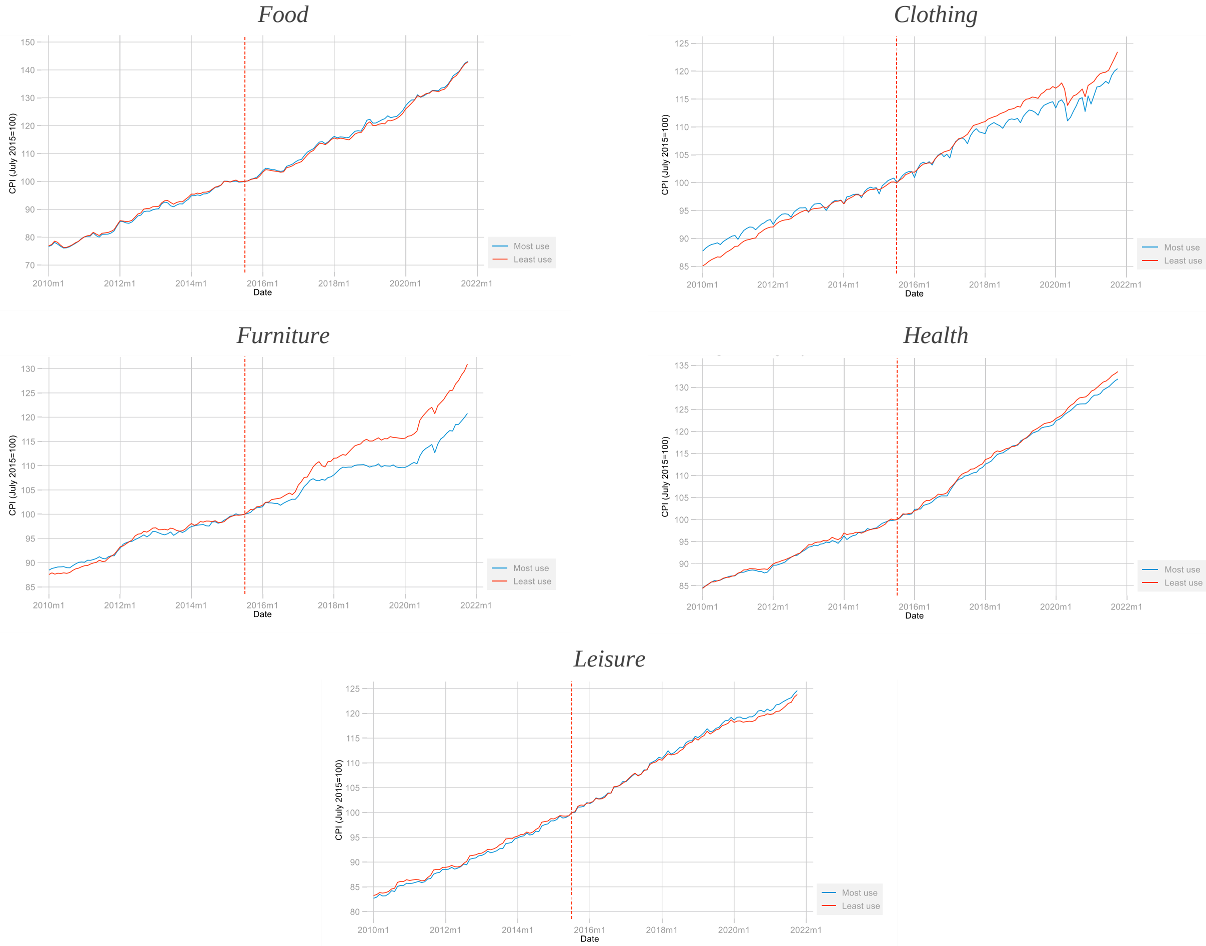


Note: Authors' calculations. The vertical dotted line indicates the date on which Amazon entered Mexico.

We compare the states with the greatest number of internet consumers in proportion to population (ranked 14th or above) against those with the lowest number (ranked 15th or below). First, we analyze the general CPI of the cities with high levels of e-commerce against those with low levels. As Figure 1 shows, there is no visible significant difference between the aggregate of cities with high levels of e-commerce versus those with low levels.[12] This is expected, as many products considered basic by INEGI (and highly weighted in the CPI index aggregation) are still not easily sold online by Amazon, and the CPI also includes services like housing rent and transportation. This is a highly aggregated measure that confirms that the general CPIs of both groups of cities show a similar general trend. This will be useful later, when we estimate a difference-in-differences effect, and it can confirm that the parallel trends assumption, which is needed to determine causality, holds: both treatment and control groups behave similarly in the aggregate.

[12] We consider the date of Amazon's entry into Mexico to be July 2015, as this was when it started selling products other than books and Kindles.

Figure 2. Weighted Average of City CPIs, by E-commerce Customers per Capita, by Product Group.



Notes: Calculations made by authors. The vertical dotted line indicates the date on which Amazon entered Mexico.

ENDUTIH has the advantage of providing information on the products that were bought online. There are some interesting observations, for example that Yucatán has a remarkably large number of people who order food online (both in absolute terms and per capita) that do not buy any other type of goods via the internet. We can recalculate the same CPI Figures, but by product group. As shown in Figure 2, there is an observable effect for clothing, furniture, and health goods: cities with the highest proportion of e-commerce users tend to have lower levels of price growth in those goods following Amazon's entry into Mexico. There is no effect for food. The curious exception is for leisure, which shows an opposite trend, but the difference is not large and may not be statistically significant.[13]

It should be noted that the different CPIs showed similar trends before Amazon's entry, which provides support for the idea that the lower observed prices (particularly in furniture) are driven by Amazon. There are some generalizations in the analysis that should be pointed out:

1) The number of users can only be obtained at the state level, and the CPIs are only reported by city. Given our data limitations, we cannot clearly distinguish whether a particular city could have a lower or higher level of e-commerce users than its state. However, given that rural residents have little access to internet

[13] The "leisure" CPIs include leisure and education. These are mainly services (private education, entertainment services, newspapers, magazines, cinema), but they also include some entertainment goods, including photographic equipment, musical instruments, and electronic and board games. The latter goods are the only items that could have been affected by competition from Amazon, as will be explained below.

services we believe there is a high correlation between e-commerce shares of users in states and in cities within the states. Given our method of dividing states with high and low levels of e-commerce, it would be unlikely for there to be a state with a high level of e-commerce that had cities with a low level.

2) We consider the date of Amazon's entry into Mexico to be July 2015, as this was when Amazon started selling products other than books and Kindles. However, the exact entry of every type of product is not considered (for example, Amazon started selling food products only recently).
3) There may be some confounding effects here, particularly since Amazon may not be the only e-commerce retailer that entered Mexico during this period. This concern will be addressed below, as we also analyze microdata on a product-by-product basis, and we control for multiple fixed effects. With these microdata, we should be able to convincingly confirm or reject our hypothesis, as we will be able to determine the exact date on which Amazon started selling each type of product.

To determine the statistical significance of these results, we apply a simple difference-in-differences (DID) analysis to the data, as described in the following sections.[14]

2.2. Estimating DID Effects on the CPI

Using a simple difference-in-differences (DID) methodology, we compare the differences in the weighted average CPIs of cities with higher and lower use of e-commerce, before and after Amazon's entry.[15] DID models allow for the control of omitted fixed variables, that is, factors that are invariant over time. The method identifies a group of variables (products, markets, or companies) affected by an event, and a second group not affected by it. The first group is the treatment group and the second is the control group. The analysis allows the temporal comparison of the market before and after the treatment with respect to a comparison market. The key identifying assumption with this method is that the trends for the treatment and control groups before Amazon's entry are similar. The DID analysis starts with the estimation of the following equation:

$$CPI_{i,t} = \beta_0 + \beta_1 \cdot tperiod_t + \beta_2 \cdot treated_i + \beta_3 \cdot tperiod_t \cdot treated_i + \varepsilon_{i,t},$$

where

- $tperiod_t$ is a dummy variable that takes the value 1 in the follow-up period (from July 2015), after the entry of Amazon (note that it does not vary between populations) and 0 in the baseline (before treatment);
- $treated_i$ is a dummy variable that takes the value 1 in the city *i* that was affected or treated (note that it does not vary over time; the treatment population will be the CPIs of the cities with the highest proportion of e-commerce consumers, per product group); and
- $tperiod_t \cdot treated_i$ is the interaction of these dummies.

The estimated coefficient of interest is $\widehat{\beta_3}$, which measures the DID effect. The period of analysis is January 2010 to December 2020; the dividing line is at the midpoint of the time series, with the same number of observations before and after.

[14] The results here are in line with the work of Cavallo (2017), who found that online and offline price levels tend to have strong similarities and converge over time. This could be the reason for the more profound price effect on brick-and-mortar prices in cities with a higher degree of e-commerce penetration. In the next section we corroborate this effect.

[15] Recall from the previous section that our method of dividing cities into those with low and high levels of e-commerce controls for the fact that the growth rates of e-commerce could be different in some places than in others, as most were in the same group in 2015 as in 2020. This is much more important if we look at e-commerce users in absolute terms, so a sensitivity analysis in those terms is also described in Appendix A. It obtains similar results.

Table 4. DID Coefficient of the Regressions on CPI.

Coefficients	(1) **Food**	(2) **Clothing**	(3) **Furniture**	(4) **Health**	(5) **Leisure**
$\widehat{\beta_3}$	1.090	-2.391**	-3.647***	-0.336	0.849
	(2.190)	(1.106)	(1.141)	(1.651)	(1.409)
$\widehat{\beta_0}$ (constant)[16]	88.472	93.499	94.364	92.422	91.214
Observations	264	264	264	264	264
R-squared	0.707	0.758	0.729	0.722	0.761

Robust standard errors in parentheses. Each column is a separate regression.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

As we can see in Table 4, the effects on leisure, health, and food are not significantly different from zero. However, the effects on clothing and furniture are negative, and these are significantly different from zero. This is an expected result, as the only two CPI indexes that include mostly goods sold by Amazon are those for clothing and furniture, and these are the only two indexes in which we find statistically significant effects. The difference is large. It means that the CPI in cities with a high share of e-commerce users is, on average, 2.4 percentage points lower for clothing and 3.7 percentage points lower for furniture. This supports the hypothesis that the effects are probably caused by the entry of Amazon: the indexes including goods and services not sold on Amazon do not show the same behavior. The causal effect is more striking given that our method for dividing cities with high and low levels of e-commerce helps to control for the fact that penetration rates might change over time, as more than 90% of them stay in the same group throughout the period analyzed.

In sum, we find that i) the general CPI between cities with high and low levels of e-commerce users moves in a parallel manner, ii) the CPI by product group diverges only for groups that include mainly products sold on Amazon (and those in which Amazon sells few are not significantly different), and iii) the trends for every group before Amazon's entry are the same between the treatment and control groups. These results are preliminary evidence for a causal effect: that Amazon's entry had a pro-competitive effect and increased price competition in Mexico. The analysis in the following section sheds additional light on the causality of this effect.

[16] This can be also interpreted as the mean of the control group, before the treatment.

3. Analysis with Microdata

3.1. Estimating the Effects of Amazon's Entry

The results from the preceding section provide important insight because they support our preliminary hypothesis that the entry of a new supplier into the retail market (in this case, Amazon) could have such an important effect on existing prices in the retail market. Given that Amazon's entry is driven only through e-commerce, its competitive discipline in the retail market could affect different regions in differing degrees, mainly depending on their development and adoption of technology.

The number of products available on Amazon has also constantly been increasing and different types of products have different times of entry. For this reason, it would be ideal to analyze their entry on a product-by-product basis to obtain an idea of the competitive discipline that brick-and-mortar establishments face each time Amazon begins to sell a product they already sell. Amazon also sells products in two different ways: "sold and delivered by Amazon" or sold by a third party, where the product fulfillment can be performed by Amazon or by the third-party seller itself.[17] Though the three options might be similar or the same from the customers' point of view (except perhaps the benefits of an Amazon Prime subscription for products sold by Amazon), there could be important differences for the suppliers selling their products on the Amazon platform.[18]

There are thus many effects in play, including efficiencies in logistics, Amazon fees to sellers, and storage costs. For these reasons, and given our data limitations, we carry out three analyses:

i) We analyze the entry of products "sold and fulfilled by Amazon." These products take full advantage of the vertical integration and the efficiencies in logistics.

ii) We analyze the entry of products sold by third parties, which may be fulfilled either by Amazon or by the third party. It would be ideal to separate these, but the Keepa data is not disaggregated to this level. Third parties might take advantage of Amazon efficiencies only in the last mile or take advantage of their logistics chains (not all third parties are small sellers; some are large companies and could have their own last-mile logistics).

iii) We analyze only the entry of products, irrespective of the seller or fulfiller. We consider only that a product started to be sold on the Amazon platform and how this entry could pose a competitive threat and affect the prices of traditional brick-and-mortar retailers.

As we do not have an a priori assumption regarding which of these alternatives have a larger competitive effect on the retail market, we analyze all three. We aim to determine, in all of the three cases, whether the entry of a product on Amazon meant competition with respect to that product in traditional brick-and-mortar retail, as we found in the preceding CPI analysis. To explore the hypothesis more deeply, we analyze microdata on average brick-and-mortar prices and Amazon products. For this analysis, we keep only the 500 products with the most balanced data in the INEGI average price dataset and those for which a similar product could be found on Amazon. As this is a novel database that we created, we describe it further in Appendix B. We implement an event study analysis (see a similar example in Atkin, Faber, and Gonzalez-Navarro 2018).[19]

Given that this microdata has much more detail, up to the product and monthly level, we propose an event study design that allows us to transparently and nonparametrically test whether Amazon influenced prices or if there was a preexisting price trend. This data also allows us to analyze, by product, those groups that were

[17] There are actually three options, sold and fulfilled by Amazon, sold by a third-party and fulfilled by Amazon, and sold and fulfilled by a third-party. The microdata from the Keepa API shows the information in two groups, by type of seller: Amazon (sold and fulfilled by Amazon) and third-party (sold by a third party, either fulfilled by Amazon or not). The data is described in more detail below.

[18] On the different logistics models, see: https://vender.amazon.com.mx/enviar?ref_=sdmx_soa_pymes_fulfill (Spanish version).

[19] Our data is not as rich as that in Atkin, Faber, and Gonzalez-Navarro (2018). They have extraordinarily rich microdata on CPI products, on an almost perfectly balanced set of products (to the municipality level). We use average prices to the product level, by city as reported by INEGI. We do not observe store prices.

discarded for lack of effect in the CPI analysis, because the CPI includes many products and services not available on Amazon; with this microdata, however, we can observe only those products found on the platform.

Estimating the treatment effect before the entry of a product allows us to test for the presence of trends in the run-up to its entry without imposing parametric structure, but still observing the detail on the slope and trend of the effect. This analysis gives better insight than a simple before/after effect, because we can see its monthly evolution and whether it is increasing or decreasing over time.[20]

We use our microdata to estimate the following baseline event study specification:

$$\ln p_{b,g,c,t} = \sum_{j=-13}^{49} \beta_j I(\text{Months since Entry}_{b,t} = j) + \delta_g + \gamma_c + \eta_{c,t} + \epsilon_g$$

where $\ln p_{b,g,c,t}$ is the log price of product b in product group g, in city c and month t; $I(\text{Months since Entry}_{b,t} = j)$ is an indicator function; and Months since Entry counts the months since the product entry for each product b at time t (with negative values counting months before entry, positive values counting months after entry, and zero being the month the product was initially listed on Amazon).[21] The β_j parameters capture the effect of Amazon or third-party entry on domestic prices for each of j months before and after the entry of the product, δ_g is a by-group (to the "generic" level) fixed effect, γ_c is a by-city fixed effect, and $\eta_{c,t}$ stands for city-date (year and month, per city) fixed effects. The dummies before the beginning of the treatment help us to observe the effects before the entry of the product on Amazon. This will help us to determine whether the effects found began with the product entry or if a preexisting downward trend could be causing the results found.

Other considerations:

1) The first and last dummies are important, as they represent the trends before and after the event analysis. They should therefore be around zero. The dummy equal to -13 represents 13 months before the effect and before, and the dummy equal to 49 represents 49 months after the effect and after, in order to aggregate the before and after effects.
2) Preexisting trends are absent in almost all groups. However, there are preexisting trends in clothing that cannot be disregarded. We are not able to eliminate these trends because of the nature of the data: a single type of clothing as followed by INEGI becomes hundreds of different products in Amazon (the same item of clothing but with many different models, sizes, and colors). This means that a single product could be matched to many different dates of entry, which creates a large variance in the month of entry. We therefore omit clothing from the analysis.
3) There are three regressions. In the first, the date 0 is when an Amazon product enters the market; in the second, the date 0 is when a third-party seller starts selling the product on Amazon; and in the third, the date 0 is when either Amazon or a third-party seller started selling the product on Amazon (whichever was first).
4) The base level is set to -1.
5) We have the weight for each group of products and city, as defined by INEGI. We therefore perform a weighted regression to give more weight to the products that contribute more to the CPI.

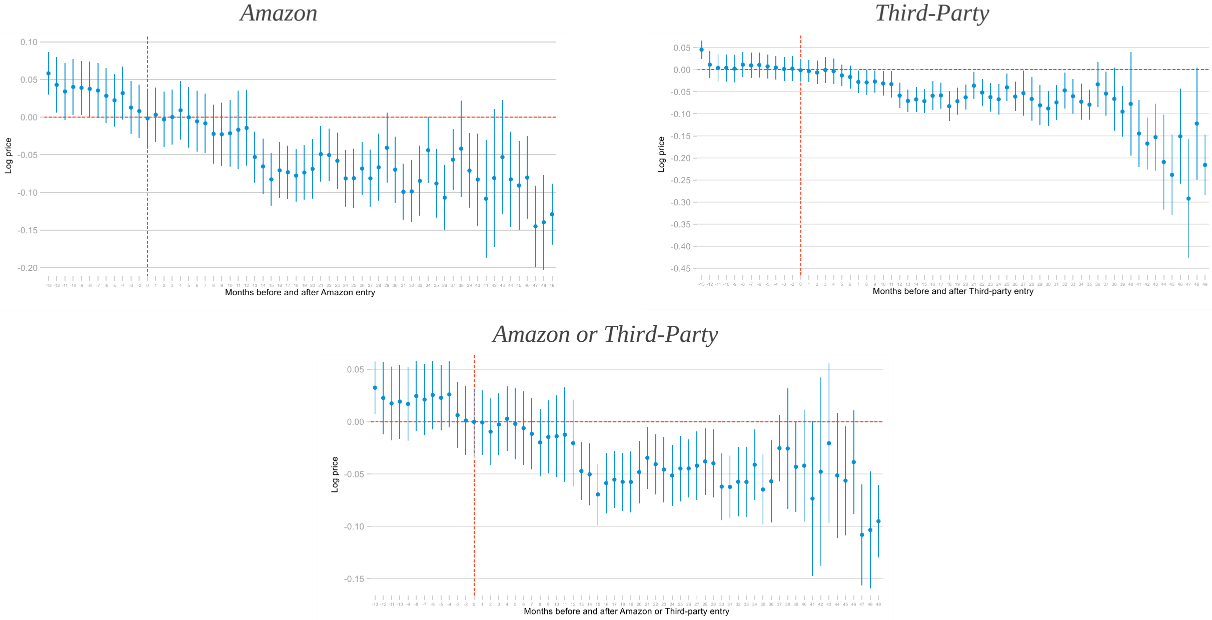
The results are shown in Figure 3. The advantages of the event analysis are clear in these graphs, where we plot the coefficients of each month-specific dummy (β_j). With these figures we can see the slope and trend of

[20] Recall that Atkin, Faber, and Gonzalez-Navarro (2018) found that the estimated gains were specific to foreign entry, rather than being driven by the entry of modern store formats more generally; their event analysis controlled for fixed effects of municipality-by-barcode-by-month. We emulate their analysis and control for every possible fixed effect that our data permit.

[21] We take $j = -1$ as the (omitted) reference category and define the indicator variable to take the value 1 for all $j \geq 49$, and similarly to take the value 1 for all $j \leq -13$.

changes and whether there was a preexisting trend or if the trend began when the product began to be sold on Amazon (β_0 vertical red line). We can see if the effect is persistent over time or just a temporary shock, and we can observe whether the effect was immediate or gradual.

Figure 3. Coefficients of Fixed Effect Weighted Regression. All products, Monthly.



Notes: Authors' calculations based on the microdata. This is the plot of the coefficients estimated: the event dummies, with 95% confidence intervals shown. It includes leisure but does not include clothing.

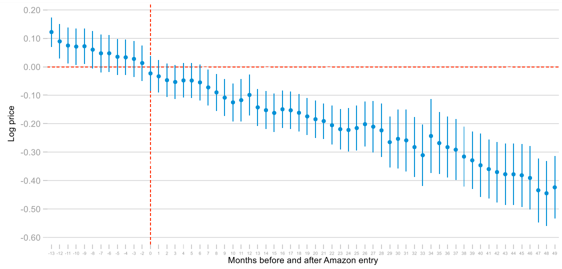
In general, there is a slight downward effect on prices, particularly after the first year of entry. This is around 5%, and it is statistically significant. We observe a similar pattern whether it is the entry of an Amazon or third-party product, or both.

We also calculate a yearly average, shown in Appendix C.[22] In the data analyzed at the yearly level, we observe the same trends.[23] For a deeper analysis, we analyze this effect with the regression on CPI groups.[24] The results are shown in Figure 4.

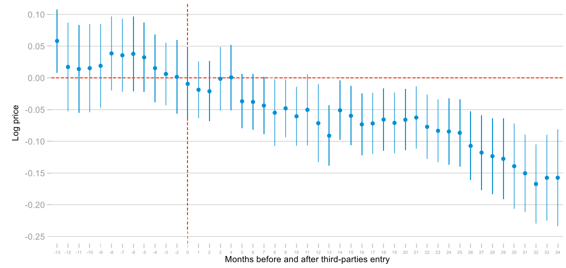
[22] The results of the regressions, both monthly and yearly, are shown in Appendix D.
[23] The advantage of the yearly analysis is that the results are less noisy. The disadvantage is that data aggregation obscures the dynamic effects of Amazon's entry. The analysis also works as a sensitivity analysis to determine whether the effects found hold in the long term.
[24] Leisure cannot be analyzed by itself as there are too few observations.

Figure 4. Coefficients of Fixed Effect Weighted Regression. Monthly. 1. Food, Beverages, and Tobacco.

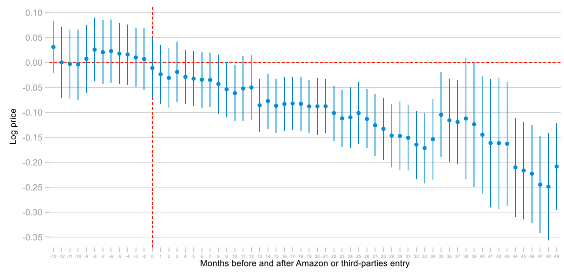
Amazon



Third-Party



Amazon or Third-Party

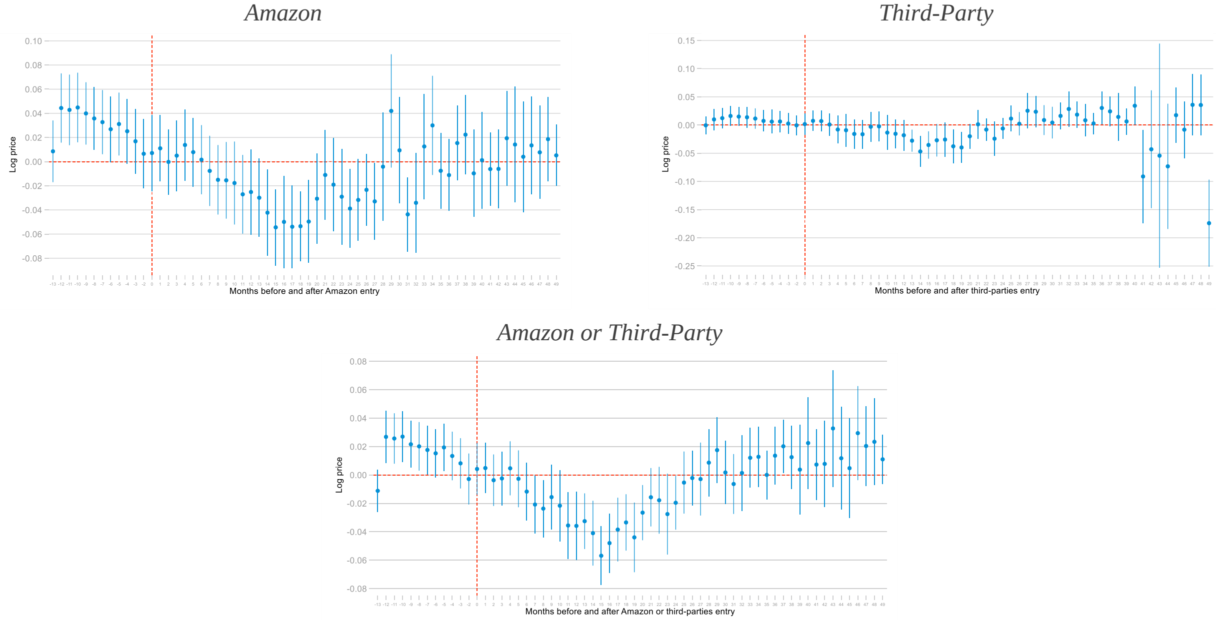


Notes: Authors' calculations based on the microdata. This is the plot of the estimated coefficients: the event dummies, with 95% confidence intervals shown.

Figure 4 shows a large downward effect on food, beverages, and tobacco, which seems to be persistent and significant, and keeps growing. This effect is approximately 10% a year for Amazon products only, 5% a year for third parties, and 7% for either. The only concern, in this case, is that there seems to be a preexisting downward trend in Amazon products only and third parties only, as the first dummy that absorbs the effect of previous months is greater than zero and significant. It is, however, non-significant in the estimation for Amazon or third-party, which suggests that the reduction in prices was probably caused by Amazon's entry.

There was no effect found in some categories, including food, in the CPI analysis. As explained in that section, the CPI aggregates various products and services. It includes some products sold by Amazon, but because most are not, we found no significant effect. With this microdata analysis, however, we can show that Amazon did have an important and significant effect on some of the few products in this category that are available on Amazon. In the case of food, this could be canned or processed food.

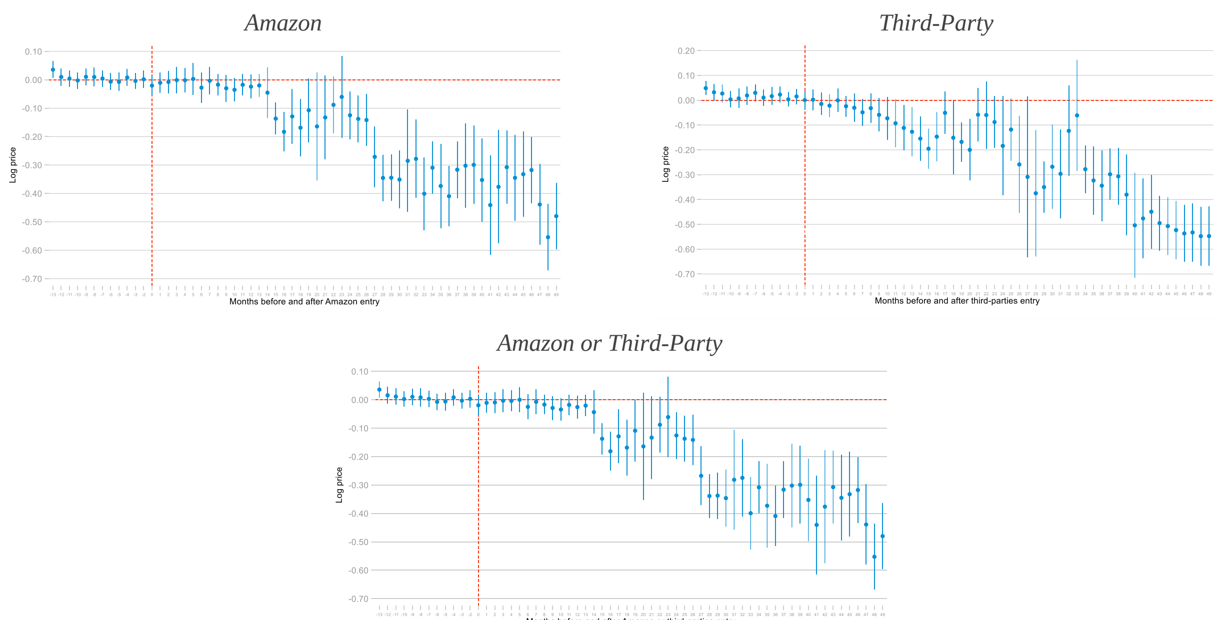
Figure 5. Coefficients of Fixed Effect Weighted Regression. Monthly. 4. Furniture, Appliances, and Household Accessories.



Notes: Authors' calculations, based on the microdata. This is the plot of the estimated coefficients: the event dummies, with 95% confidence intervals.

As can be seen in Figure 5, there is a large and significant downward effect on furniture prices, but it seems transitory. Its magnitude is approximately 7% for the first two years, but it vanishes in the analysis for sales by Amazon, and it is approximately 3% in the analysis for Amazon or third-party sales. There is no statistically significant effect on furniture for third-party sales only.

Figure 6. Coefficients of Fixed Effect Weighted Regression. Monthly. 5. Health and Personal Care.



Notes: Authors' calculations, based on the microdata. This is the plot of the estimated coefficients: the event dummies, with 95% confidence intervals.

As seen in Figure 6, there is a very large and significant downward effect on the prices of health products. The effect is null in the first year, but after that it increases. It is approximately 10% a year after Amazon entry for third-party sales or for Amazon or third-party sales. This effect increases over time, however, and reaches levels of approximately 30% four years after the availability of these products on the platform, which could mean that competitive pressure increases over time, considering that the adoption of e-commerce is also increasing over time.

We do not report an analysis for leisure, as there are too few observations for this group. However, the few products in this category are included in the aggregate results for all products, in Figure 3. Our yearly analysis by groups is also reported in Appendix C, and the results are qualitatively similar.

Table 5 summarizes the results. It shows that the entry effect of products sold and delivered by Amazon is strong and statistically significant, ranging from a downward effect on prices of 1% to 28%. The entry effect of products sold by third parties is not as strong, and is not always statistically significant, especially aggregated at the yearly level. It ranges from a 1% to 7% effect on prices, except for the 20% effect on health products.

Table 5. Summary Results of the event analysis.

	Monthly			Yearly		
	Amazon	**Third-Party**	**Either**	**Amazon**	**Third-Party**	**Either**
General	-6%*	-7%*	-4%*	-5%*	-1%	-4%
Food	-20%*	-7%*	-10%*	-28%*	-4%	-24%*
Furniture	-1%*	-1%	-1%*	1%	-5%	1%
Health	-18%*	-20%*	-18%*	-23%*	-5%	-23%*

* Statistically significant (95%). The data reported here are the simple average effects of the relevant coefficients from the regressions above. The relevant coefficients are those from the strictly positive dummies, as those capture the effect of Amazon's entry. These are converted from natural logarithms to percentages, as the dependent variable is ln(price).

The following conclusions can be drawn from the regression results:

1) For competing products introduced by Amazon:
    a. In general, a slight downward effect is observed on prices, particularly after the first year. This effect is approximately 5% and is statistically significant.
    b. In the yearly analysis, the effect is approximately 5% for three years, but is barely statistically significant.
    c. By product group:
        i. There is a large downward effect, approximately 10% per year and increasing, on the prices of food, beverages, and tobacco, which seems to be persistent and significant.
        ii. There is a large and significant downward effect on furniture, but it seems transitory. It is approximately 7% in the first two years, but then it vanishes.
        iii. There is a very large and significant downward effect on the prices of health products. It is null in the first year, but then increases to 10% a year and is growing.
    d. By product group, yearly:
        i. There is an important statistically significant downward effect on the prices of food, beverages, and tobacco, approximately 15% a year.
        ii. There is no effect on the price of furniture.
        iii. There is a very large and significant downward effect on the prices of health products, approximately 20% a year.
2) For competing products introduced by a third-party vendor on the Amazon website:
    a. In general, an important persistent downward effect is observed on prices, particularly after the first six months. This change is statistically significant. It begins at 5% but grows over time.
    b. In the yearly analysis, there is a downward effect on prices, but it is not statistically significant.
    c. By product group:

        i. There is an important downward effect of approximately 5% on the prices of food, beverages, and tobacco, which seems to be persistent, significant, and growing.
        ii. There is no statistically significant effect on the price of furniture.
        iii. There is a very large and significant downward effect on the prices of health products, approximately 5% a year and growing.
    d. By product group, yearly:
        i. There is a downward effect on the prices of food, beverages, and tobacco, but it is not statistically significant.
        ii. There is no effect on the price of furniture.
        iii. There is an important effect on the prices of health products, though it is not statistically significant.
3) For competing products introduced either by Amazon or a third-party vendor on the Amazon website:
    a. In general, a slight downward effect on prices is observed, particularly after the first year. The next three years show an effect of 5%, which then becomes statistically non-significant.
    b. In the yearly analysis, there is a non-statistically significant downward effect of approximately 7% for four years.
    c. By product group:
        i. There is an important, persistent, and significant downward effect of approximately 10% per year on the prices of food, beverages, and tobacco, throughout the whole study period.
        ii. There is a significant downward effect on furniture, but it seems transitory. It is approximately 3% for the first two years, but then it vanishes.
        iii. There is a very large and significant downward effect on the prices of health products. It is null in the first year, but then reaches 10% a year and grows.
    d. By product group, yearly:
        i. There is a statistically significant downward effect, approximately 10% per year, on the prices of food, beverages, and tobacco.
        ii. There is no significant effect on furniture.
        iii. There is a very large and significant downward effect on the prices of health products, around 15% a year.

In the next section, we will replicate this analysis with the INEGI-Keepa microdata to verify that the observed decrease in prices is due to Amazon's entry. There should be a difference in the effect on prices depending on the degree of Amazon's penetration: the greater the number of e-commerce buyers, the greater the degree of competition that Amazon's entry could pose to brick-and-mortar products.

3.2. Estimating the DID Effects with Microdata

We replicate the DID results taking advantage of the INEGI-Keepa microdata. Given that this microdata provides much greater detail, up to the product and monthly levels, we can analyze only products with a corresponding product on the Amazon platform, rather than an aggregate bundle of products, some of which may not be affected by Amazon at all.

This data also allows us to analyze, at the product level, those groups with no effect in the CPI analysis, groups that were discarded because the CPI aggregates many products and services not available on Amazon. With this microdata we can observe only those products found on the platform. In this case, we can also control for fixed effects, given the disaggregation of the microdata.

Using a simple DID implementation, we compare the differences among the weighted brick-and-mortar prices (as reported by INEGI) of cities with higher and lower levels of e-commerce, before and after Amazon entry, at the product level.[25] We consider the use of e-commerce as the percentage of people who bought products using e-commerce with respect to the total population.

The DID analysis begins with the estimation of the following equation:

$$\ln(price)_{b,g,c,t} = \beta_0 + \beta_1 \cdot tperiod_{t,b} + \beta_2 \cdot treated_{c,b} + \\ + \beta_3 \cdot tperiod_{t,b} \cdot treated_{c,b} + FE_t + \varepsilon_{i,t},$$

where

- $\ln\,(price)_{b,g,c,t}$ is the natural logarithm of the price of product b in product group g, in city c, and month t;
- $tperiod_{t,b}$ is a dummy variable that takes the value 1 in the period (beginning in July 2015; the entry date of each period is different) after the entry of Amazon (note that it does not vary between populations, except that the exact date is different for every product) and 0 in the baseline period (before treatment);
- $treated_{c,b}$ is a dummy variable that takes the value 1 in the affected population (note that it does not vary over time; the treated population is the cities *c* with the highest share of e-commerce consumers, per product group);
- $tperiod_{t,b} \cdot treated_{c,b}$ is the interaction of the dummies;[26]
- in this case, we also add FE_t, which are various dummies to control for fixed effects by date (January 2019 = 1, February 2019 = 1, and so forth).

Since the entry date is different for every product, we homogenize the index as in the previous section. Thus, the date index is equal to 0 when the product enters Amazon (either sold by Amazon, a third party, or either one). We calculate robust standard errors and use INEGI weights on the city-group of products, as before.

The estimated coefficient of interest is $\widehat{\beta_3}$, which measures the DID impact. The period of analysis ranges from -12 to 24. That is, we compare the average prices from one year before the relevant entry (Amazon, third-party, or either) to two years after the entry. We restrict our analysis to these indexes only (-12 to 24) to keep a balanced group of observations, as only a few products have information after 24 or before -12.

As in the previous DID, treated populations are the cities that, according to ENDUTIH, have a greater proportion of e-commerce buyers with respect to their population (at the state level, a rank of 14th or above). The control populations are the other cities followed by INEGI (with low levels of e-commerce users). We report the results with a relative measurement (e-commerce users with respect to the total population), and in Appendix A we briefly report the results in terms of absolute numbers of e-commerce users. The conclusions are qualitatively identical.

The treatment period will be from 0 to 24, that is, from the month of Amazon and/or third-party entry by product to 24 months later. We also estimate the effects during the 12 months before entry (-12 to -1), so that we

[25] As this analysis is at the product level, and there are different dates of entry for each competing product on Amazon, as before, instead of using the date, we re-index the data. Index = 0 when the product began to be sold on Amazon (1, 2, 3, … are one, two or three months after the entry, and -1, -2, -3, … are the corresponding months before that entry). The entry of a product is taken as the time at which a product is sold and delivered by Amazon or a third party, whichever was first.

[26] As explained above, cities' have different e-commerce adoption growth rates, so their rankings could change over time. However, treated and control groups do not change over time in our analysis given because of the way our analysis divides cities with high versus low levels of e-commerce. There is thus no issue with taking a 2020 snapshot of the share (or number) of e-commerce users, because over time more than 90% of them remain in the same group.

can visualize whether the trend we find is a preexisting one or is due to the entry of products onto the Amazon platform.

Table 6. DID Results for All Products

	All products		
	(1)	(2)	(3)
Coefficients	Amazon	Third-Party	Either
$\widehat{\beta_3}$	-0.366***	-0.449***	-0.449***
	(0.0485)	(0.0455)	(0.0419)
Observations	22,886	29,080	34,099
R-squared	0.075	0.116	0.047

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

Table 6 shows that the downward effects of Amazon on prices are significant and much more important in cities with higher levels of e-commerce users. Prices in these cities decline by 31% (obtained from 1-exp (-0.366)). We carry out this same analysis by product group.

Table 7. DID Results for Product Groups

	Food			Furniture			Health		
	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Coefficients	Amazon	Third-Party	Either	Amazon	Third-Party	Either	Amazon	Third-Party	Either
$\widehat{\beta_3}$	-0.0158	-0.375***	-0.132**	-0.364***	-0.237***	-0.295***	-0.224***	0.0566	-0.220***
	(0.0750)	(0.0550)	(0.0533)	(0.0609)	(0.0654)	(0.0643)	(0.0646)	(0.0654)	(0.0609)
Observations	10,398	16,601	19,871	10,726	10,596	12,245	1,716	1,696	1,786
R-squared	0.078	0.084	0.032	0.118	0.264	0.069	0.486	0.477	0.457

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

In Table 7 we also observe that Amazon's downward effects on prices are significant and much more important in cities with higher levels of e-commerce users, for almost any product group. The CPI analysis found no effect on food or health; the CPI aggregates many more products and services, most of which are not sold on Amazon. The analysis in this section uses microdata including only a subset of products within the aggregated groups, all of which have a closely corresponding product on the Amazon platform. The results are summarized in Table 8.

Table 8. Summary of DID Results

	DID Results		
	Amazon	**Third-Party**	**Either**
General	-31%***	-36%***	-36%***
Food	-2%	-31%***	-12%**
Furniture	-31%***	-21%***	-26%***
Health	-20%***	6%	-20%***
Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$			

As can be seen in Table 8, there are large and statistically significant effects. Moreover, these are large and statistically significant for the "general" category, which includes all of the products (food, furniture, health, and leisure), weighted by the respective group-city weight that INEGI uses to calculate the national CPI. The third-party effect is larger than that of products sold and fulfilled by Amazon for the food and general categories. The opposite is the case for the furniture and health categories.

To provide a more intuitive view of the results, the first column in Table 8 shows the effect of the products sold and fulfilled by Amazon, the second column shows those sold by third parties, and the third column shows the effect of the entry of the product onto the platform, without regard to how it was sold or fulfilled. The third column can be seen as a weighted average of the other two columns. We can use this information to differentiate the effect of entry from fulfillment-specific effects. For example, if we include all products (the “general” category), we can see that “either” has the same effect as “third-party”: that is, the stronger price reduction effect. Almost all of the downward effect on brick-and-mortar prices is caused by the entry of the products per se. The pro-competitive effect of the entry of a new product is the same, whether it is sold by Amazon or by a third party using the Amazon platform.

The column “either” shows the main effects of Amazon competition, because it indicates price movement after the first entry of a product, whether it was sold by Amazon or a third-party seller. The effects on health and food products, which are large and statistically significant, come mainly from products sold by Amazon. There are also statistically significant effects on furniture, either sold by Amazon or by third-party sellers. We can therefore conclude that the downward price effects of Amazon are much more important in cities with high levels of e-commerce users, which supports the hypothesis that the decrease in prices is caused by the entry of Amazon.

4. Preliminary Welfare Estimations

We find a significant decrease in prices after Amazon's entry. At the CPI level, Amazon's downward effects on brick-and-mortar prices are observed only in clothing and furniture, which are the CPI categories that include many products (but not services) sold by Amazon and in which Amazon could pose competitive pressure. At the microdata level, we find an Amazon effect also in the prices of food and probably also leisure (though we cannot be sure because there are not enough products in leisure to analyze that group by itself). However, because Amazon does not sell many products in these CPI categories (which include raw perishable food, services, and other types of goods), the effects might be less obvious.

At the CPI level, we find significant effects in clothing and furniture. We thus conservatively estimate an effect, assuming that these were the only two product groups in which Amazon has an effect, because the CPI decreases in a statistically significant manner. We emulate the first-order approach of welfare estimation using observed price differences (see Atkin, Faber, & Gonzalez-Navarro 2018, pp. 13-14). We thus estimate the cost-of-living effect by approximating the expenditure function through a first-order Taylor approximation: as the welfare change due to the price changes induced by the arrival of Amazon. We decompose the cost-of-living effect into the direct price index effect (the effect of Amazon's low prices) and the pro-competitive effect (the effect of brick-and-mortar price decreases due to Amazon).

However, Atkin, Faber, and Gonzalez-Navarro (2018) exploit the store price data to estimate a first-order approximation of the cost-of-living effect based only on observable price changes. As we do not have this level of microdata disaggregation, our analysis must be on the aggregate. They show that a first-order approximation of the pro-competitive effect is a Paasche price index of the product-level price changes at continuing incumbent stores due to entry, multiplied by the period 1 share of total expenditure captured by that store-product pair. Nevertheless, this explicitly assumes no exit of stores (pro-competitive exit of less efficient suppliers given the strong competition brought by the new more efficient entry). We will have to assume no exit for this period.

We take a first-order Taylor expansion of the expenditure function around last period prices and apply Shephard's lemma. By focusing on sales and price changes in the set of stores continuously selling a particular product across periods 0 and 1 (for which we can observe price changes), we obtain the pro-competitive price effect. For the direct price index effect, we focus on the sales and price changes in the Taylor expansion around prices in period 1, which we assume to be around 2020.[27,28]

The equations to be calculated, with our information on the aggregate, are

$$pce \approx \sum_{b} \sum_{s \in S_{b,ecom}^{hec}} \left(\varphi_{b,sh}^{1} \left(CPI_{b,lec}^{1} - CPI_{b,hec}^{1} \right) \right)$$

for the pro-competitive effect (pce), and

$$dpie \approx \sum_{b} \sum_{s \in S_{b,bnm}^{hec}} \left(\varphi_{b,sh}^{1} \left(CPI_{b,lec}^{1} - CPI_{b,hec}^{1} \right) \right)$$

for the direct price index effect (dpie), where $\varphi_{b,sh}^{1}$ is the household expenditure share spent on the product b (either clothing or furniture) in period 1 (2020), S_b^{hec} is the set of cities that have high levels of e-commerce users

[27] For simplicity, in contrast to the method of Atkin, Faber, and Gonzalez-Navarro (2018), we assume the entry as 2020 for all products.
[28] In Atkin, Faber, and Gonzalez-Navarro (2018) the direct price index effect corresponds to a Paasche price index of the product-level price differences between foreign stores in period 1 and domestic stores in period 0, multiplied by the period 1 share of total expenditure captured by foreign stores for that product.

of product b (clothing or furniture) in period 1 (either be for *ecom*, e-commerce, or *bnm*, brick-and-mortar), and $CPI_{b,lec}^{1} - CPI_{b,hec}^{1}$ is the difference between CPIs (in 2020) of low- and high-level e-commerce consumers, for product b (clothing or furniture) in period 1.

As we do not have access to their remarkable level of microdata, we must replicate their calculation but at an aggregate level. We calculate the direct price index effect and the pro-competitive effect using:

- the share of clothing and furniture in expenditure (0.024 for clothing and 0.064 for furniture, on average);[29]
- multiplied by the Amazon post-entry market share (0.050 for clothing and 0.047 for furniture, on average) for the direct effect,[30] and the brick-and-mortar post-entry market share (0.95 for clothing and 0.953 for furniture, on average) for the pro-competitive effect;
- multiplied by the CPI post-entry gap between high and low levels of e-commerce consumers (0.025 for clothing and 0.070 for furniture, on average, in 2020).[31]

This means we have to make several additional assumptions:

- We obtain only the effects on the 14 states with the greatest proportion of e-commerce consumers, given that we find that the cities in those states were the ones with lower CPIs. The effect is thus strongly underestimated, as the price decreases also occurred in the states with low levels of e-commerce users, but to a lesser degree. We will, however, maintain this assumption to keep our welfare estimates as conservative as possible. The underlying assumption is thus that there is no effect in cities with low levels of e-commerce.
- We may be underestimating the share of expenditure in Amazon, as we analyze only clothing and furniture, not retail as a whole. Amazon also participates to a lesser extent in food and leisure, but given the small size of this participation, we could not find a statistically significant effect at the CPI level.
- We may be underestimating Amazon's 2020 market share, as our assumption regarding the percentage of use of every customer is too high[32] (this means that some of the pro-competitive effect could be from the direct price index effect).
- We are also underestimating the actual Amazon effect on prices, as our best information comes from the CPIs, and these are highly aggregated measures.
- We do not have enough data to calculate the other effects (more product variety, for example), which have been shown to have a great impact on the welfare effect (Atkin, Faber, & Gonzalez-Navarro 2018).

Table 9. Direct Price Index and Pro-competitive Effect, by Product Group

	Clothing			Furniture		
	Total Effect	**Direct Price Index Effect**	**Pro-competitive Effect**	**Total Effect**	**Direct Price Index Effect**	**Pro-competitive Effect**
Average Effect	0.0006*** (0.000008)	0.0000295*** (0.0000004)	0.00057*** (0.000007)	0.00445*** (0.00003)	0.0002111*** (0.0000016)	0.00424*** (0.00003)

[29] We use ENIGH 2020 microdata (the publicly available version, in "*concentradohogar*") for "current monetary expenditure" (*gasto_mon*) for total expenditure, and expenditure on "clothing and footwear" (*vesti_calz*) and "cleaning" (*limpieza*). "Cleaning" aggregates "expenditure on articles and services for cleaning, home care, household goods and furniture, glassware, household utensils, and linen." "Current monetary expenditure" is the sum of all types of household expenditures.

[30] We approximate this with ENDUTIH 2020. As can be seen in Table 3, we use the percentage of people that made online purchases, by state. However, these people do not necessarily buy everything online. We will therefore approximate their expenditure using their reported frequency of online shopping on a weekly basis: at least once a week (1), once every fifteen days (1/2), once a month (1/4), once every six months (1/26), or once a year (1/52).

[31] We take the CPI difference between clothing and furniture in 2020. This difference is probably the closest estimate we have of the downward price effect, as it compares the prices in places with high levels of e-commerce users to those of places with low levels of e-commerce users (which we can use as a counterfactual for the evolution of prices without e-commerce in Mexico).

[32] Our assumptions imply, for example, that those who buy furniture online twice a month spend only half of their furniture expenditure online and half in brick-and-mortar establishments.

Obs. (households)	44,219	43,677

Notes: Linearized standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$
Table reports the welfare effects of Amazon entry as described in the text. The average effect is the weighted mean across all 44,000 households with high levels of e-commerce users in the ENDUTIH 2020 income and expenditure survey.

The results are shown in Table 9. Though they appear to be relatively small (0.06% of 2020 household income for clothing and 0.45% for furniture), the overall effect is extremely large, in that it affects many households (in contrast to the result in Atkin, Faber, and Gonzalez-Navarro 2018, who found a direct price index effect of approximately 0.02, with only 12,293 households). These effects have national impact, at least in the states with a larger number of e-commerce consumers. They are also increasing, as Amazon grows its logistics chain in Mexico to deliver more rapidly and efficiently and expand to additional locations in Mexico, increasing the number of e-commerce consumers.

Three are three other potentially important welfare gains related to Amazon's entry that our analysis does not capture: the availability of new product variety, the ease of buying online, and elimination of the need go to a brick-and-mortar store (particularly in the era of COVID-19), plus the additional amenity of having an extra shopping choice available at all times and with fast delivery. Atkin, Faber, and Gonzalez-Navarro (2018) found that these effects account for more than half of the total welfare gains.

5. Concluding Remarks

From the analysis of the CPIs by product group, we find that, on average, cities with more e-commerce consumers tend to have lower CPIs. This effect is large and significant in furniture and clothing, which have a substantial presence in e-commerce sales. The CPI in cities with a high proportion of e-commerce users is, on average, 2.4 percentage points lower in clothing and 3.7 percentage points lower in furniture relative to those with fewer e-commerce consumers after Amazon's entry. We find evidence that this effect is causal: Amazon generated a pro-competitive effect with a downward effect on brick-and-mortar prices, and the more e-commerce consumers the larger the effect.

We also analyze a novel database we created by merging INEGI average prices of brick-and-mortar products with information on products sold on Amazon (obtained from a third-party app, Keepa). We find that, in general, when Amazon started selling a product, brick-and-mortar prices decreased (though to different degrees depending on the type of product). The effect of new entry sales by Amazon is large and statistically significant, ranging from a 1% to 28% decrease in prices. The impact of initial sales by third parties is not as large and not always statistically significant (especially when aggregated to the yearly level), ranging from a 1% to 7% decrease in prices (except for health products, for which we found strong effects from third-party vendors).

With a difference-in-differences analysis we also find that this price effect is more important in cities with more e-commerce users. It is strong in the analysis of all products. The effects on health products come mainly from products sold by Amazon (a price decrease of approximately 20%). The price reductions in food are also generated by products "sold by Amazon" (a price decrease of approximately 31%). These effects are large and statistically significant. There are also statistically significant effects for furniture, either from Amazon or third-party sellers (price decreases of approximately 31% and 21%, respectively). The general effect is very large (a price decrease of approximately 36%).

These two statistical exercises provide strong evidence that Amazon generated a price decrease in products sold online, which could also have important pro-competitive effects and lower brick-and-mortar prices as well.

Finally, we carry out a conservative rough estimate of the welfare gains exclusively from this price decrease (disregarding other important welfare effects of e-commerce like increased product variety, increased ease of purchase, and elimination of transaction costs) and we find statistically significant gains in states with high proportions of e-commerce consumers. Our estimations are very conservative and the resulting welfare effects are relatively small (though strongly significant). The effects are at least 0.06% of 2020 household income for clothing and 0.45% for furniture. They affect Mexico on a national level, and the pro-competitive effect will increase in time as e-commerce penetration is increases in México.

References

Atkin, Faber & Gonzalez-Navarro. 2018. "Retail Globalization and Household Welfare: Evidence from Mexico." *Journal of Political Economy*, 126(1), 1-73.

Cavallo. 2017. "Are online and offline prices similar? Evidence from large multichannel retailers." *American Economic Review*, 107, 283-303.

Fajgelbaum & Khandelwal. 2014. "Measuring the unequal gains from trade." *UCLA Working Paper*.

Goldfarb & Tucker. 2019. "Digital Economies." *Journal of Economic Literature*, 57, 3-43.

He, Reimers & Shiller. 2021. "Does Amazon Exercise its Market Power? Evidence from Toys R Us." *SSRN*.

Iacovone, Javorcik, Keller & Tybout. 2015. "Supplier responses to Wal-Mart's invasion in Mexico." *Journal of International Economics*, 95(1), 1 - 15.

Javorcik & Li. 2013. "Do the biggest aisles serve a brighter future? Global retail chains and their implications for Romania." *Journal of International Economics*, 90(2), 348–363.

Palacios. 2001. "Globalization and e-commerce: Growth and impacts in Mexico."

Porto. 2006. "Using survey data to assess the distributional effects of trade policy." *Journal of International Economics*, 70(1), 140–160.

Topalova. 2010. "Factor immobility and regional impacts of trade liberalization: Evidence on poverty from India." *American Economic Journal: Applied Economics*, 2(4), 1–41.

A. CPI Trends in Absolute Terms

Data and Visual Inspection

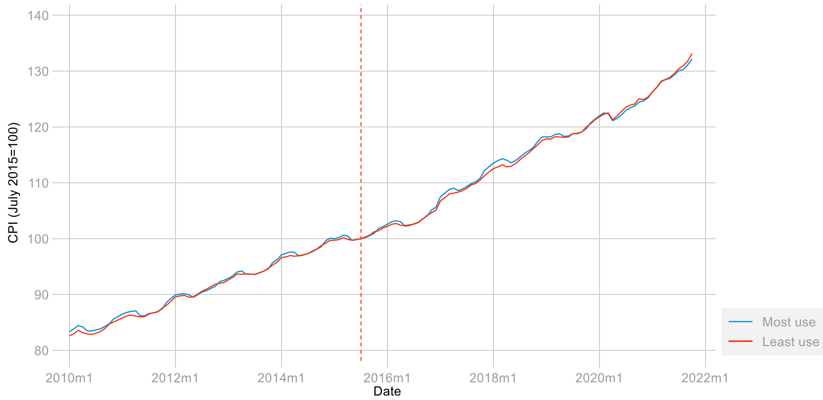
As noted in the first section, the division between areas with high and low levels of e-commerce consumers could also be made in absolute terms (that is, in terms of the absolute numbers of e-commerce consumers in each area). This division would also help to address the issue that e-commerce penetration rates in different cities may have changed over the period of analysis, but in absolute terms the cities in the groups are almost the same from 2015 to 2020. In this Appendix we briefly discuss the results in absolute terms; in short, the conclusions are identical.

Table 10. Rank of Cities in the Absolute Number of Consumers Who Use the Internet to Buy Goods

City	General	Food	Clothing	Furniture	Health	Leisure
Toluca, Edo. de Méx.	1	1	1	1	1	1
Mexico City	2	2	2	2	2	2
Guadalajara, Jal.	3	4	3	3	3	4
Tepatitlán, Jal.	3	4	3	3	3	4
Monterrey, N.L.	4	3	4	4	4	3
Córdoba, Ver.	5	7	5	5	5	7
San Andrés Tuxtla, Ver.	5	7	5	5	5	7
Veracruz, Ver.	5	7	5	5	5	7
Tijuana, B.C.	6	6	6	6	6	5
Mexicali, B.C.	6	6	6	6	6	5
Cortazar, Gto.	7	9	7	9	7	6
León, Gto.	7	9	7	9	7	6
Tampico, Tamps.	8	14	9	8	11	14
Matamoros, Tamps.	8	14	9	8	11	14
Cd. Juárez, Chih.	9	11	8	7	13	8
Chihuahua, Chih.	9	11	8	7	13	8
Cd. Jiménez, Chih.	9	11	8	7	13	8
Hermosillo, Son.	10	10	13	10	9	9
Huatabampo, Son.	10	10	13	10	9	9
Monclova, Coah.	12	13	10	12	8	12
Cd. Acuña, Coah.	12	13	10	12	8	12
Torreón, Coah.	12	13	10	12	8	12
San Luis Potosí, S.L.P.	13	18	12	13	15	16
Culiacán, Sin.	13	18	12	13	15	16
Morelia, Mich.	14	17	14	14	14	17
Jacona, Mich.	14	17	14	14	14	17
Mérida, Yuc.	15	5	17	16	12	13
Querétaro, Qro.	16	12	15	15	16	11
Tulancingo, Hgo.	18	30	18	24	24	19
Chetumal, Q.R.	19	16	20	18	18	15
Oaxaca, Oax.	20	20	23	20	20	22
Tehuantepec, Oax.	20	20	23	20	20	22
Villahermosa, Tab.	21	19	22	19	17	21
Cuernavaca, Mor.	22	21	19	22	22	24
Iguala, Gro.	23	26	21	23	21	26
Acapulco, Gro.	23	26	21	23	21	26
Tapachula, Chis.	24	24	24	25	23	20
Durango, Dgo.	25	22	25	21	26	25
La Paz, B.C.S.	27	28	30	28	29	27
Tepic, Nay.	28	29	27	29	30	31
Fresnillo, Zac.	29	31	29	27	28	28
Colima, Col.	30	27	28	30	27	29
Campeche, Camp.	31	25	32	32	31	30
Tlaxcala, Tlax.	32	32	31	31	32	32
Number of top-ranking cities	26	24	26	26	25	24
Number of bottom-ranking cities	18	20	18	18	19	20

Table 10 summarizes the ENDUTIH results in absolute terms. As expected, the highest-ranking places are Estado de México, Mexico City, Monterrey, and Guadalajara. But there are important differences for some states with low population but high penetration of e-commerce, like Colima. We compare the cities that belong to the highest-ranking states in terms of their absolute number of e-commerce users (ranked 14 or higher) against those in states with the lowest ranking (ranked 15 or lower). First, we analyze the general CPI of those cities with high levels of e-commerce users against those with low levels.

Figure 7. Weighted Average of City General CPIs, by E-commerce Customers per Capita.



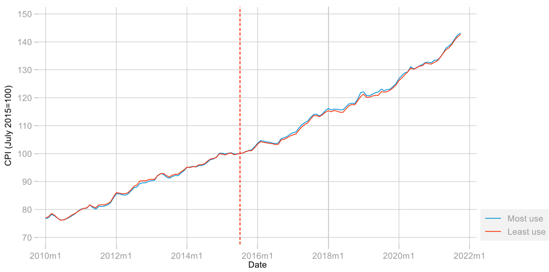
Note: Authors' calculations.

As can be seen in Figure 7, there is no visible effect in the aggregate of cities with high levels of e-commerce versus those with low levels. This was also true in the analysis in relative terms, as this is a highly aggregated measure. We can thus confirm that the general CPIs of both groups of cities have a similar general trend. This is useful for our estimation of a difference-in-differences effect, as it confirms that the parallel trends assumption that is needed to determine causality holds, as both treatment and control groups behave similarly in the aggregate.

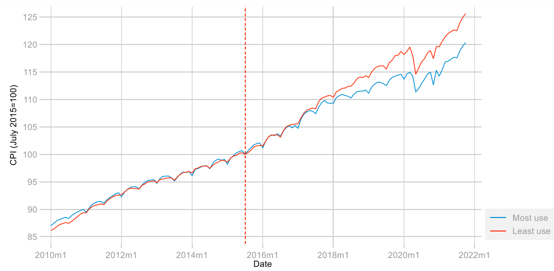
Regarding the differences by product group, Figure 8 shows the same results as the analysis in relative terms. For clothing, furniture, and health products, we observe some effect: cities with the highest number of e-commerce users (in absolute terms) tend to show lower levels of price growth in those goods after the entry of Amazon into Mexico. There is no effect for food. The curious exception is the CPI for leisure, where we observe the opposite. The difference, however, is not very large and may not be statistically significant. We therefore draw the same conclusion from this analysis as from the analysis of relative numbers.

Figure 8. Weighted Average of City CPIs, by E-commerce Customers per Capita, by Product Group

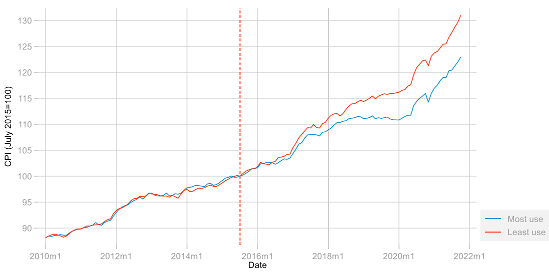
Food



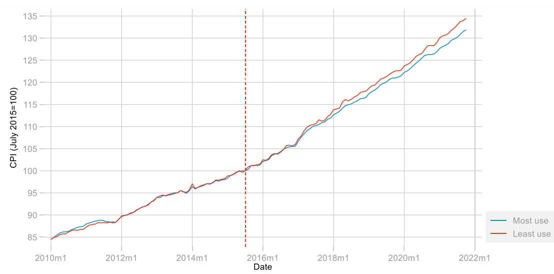
Clothing



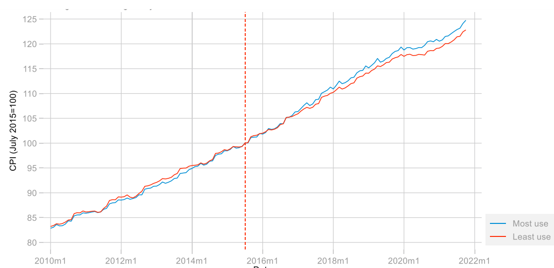
Furniture



Health



Leisure



Notes: Authors' calculations.

DID Analysis: CPIs and Microdata

We carry out a DID analysis exactly as was done with relative figures, except that the ranks of the cities are now sorted in absolute terms. As can be seen in Table 11, the effects for leisure, health, and food are not statistically different from zero. However, the effects on clothing and furniture are negative, as expected, and statistically different from zero.

Table 11. DID Coefficient of the Regressions on CPIs.

Coefficient	(1) Food	(2) Clothing	(3) Furniture	(4) Health	(5) Leisure
$\widehat{\beta_3}$	0.768	-2.043*	-2.737**	-0.978	1.262
	(2.187)	(1.140)	(1.168)	(1.672)	(1.394)
Observations	264	264	264	264	264
R-squared	0.706	0.752	0.726	0.720	0.760

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

The difference found in clothing and furniture is very important. It means that the average CPI in cities with a large number of e-commerce users is, on average, 2 percentage points lower for clothing and 2.7 percentage points lower for furniture. In the analysis in relative terms, those effects were notably larger: 2.4 percentage points lower in clothing and 3.7 percentage points lower in furniture. The results are qualitatively identical, but the magnitudes are larger when the number of e-commerce users is considered in relative terms. This supports the hypothesis that the effects, which were found to be statistically significant, are probably caused by Amazon's entry, as the indexes that include goods or services not sold on Amazon do not show the same behavior.

We also repeat the exercise of the DID with microdata. These results too are qualitatively identical, but the magnitudes are smaller when the number of cities is considered in absolute terms. For the sake of brevity, we present here only the summary table of the results (Table 12).

Table 12. DID Summary Results

	Amazon	**Third Party**	**Either**
General	-6%	-22%***	-18%***
Food	9%	-30%***	-11%**
Furniture	-5%	-4%	-6%
Health	-29%***	-2%	-27%***
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$			

There are some strong statistically significant effects, as is clear from Table 12. The main effects in all products and in food products come mainly from third-party sellers. The effects on health products come mainly from products sold by Amazon. These effects are large and statistically significant. There are, however, no statistically significant effects on furniture (though the sign is as expected). In this case, though the conclusions are the same, the resulting coefficients are notably larger (and more statistically significant) in relative terms than in absolute terms. We can therefore conclude that the downward price effects of Amazon are much more important in cities with high levels of e-commerce users, which supports the hypothesis that the observed decrease in prices is caused by the entry of Amazon. This hypothesis holds both for absolute and relative terms.

B. Data

Price Information from Amazon

The information about the prices of products on Amazon was obtained from a third-party app (Keepa), through its paid API. The process was the following:

1) From the INEGI price microdata, a list was obtained of the 500 most representative products which could be sold online (and those for which we could find information from several cities for the same product, and for a considerable length of time).[33] These brick-and-mortar prices are obtained from publicly available INEGI data.
2) These 500 products were found manually on Amazon. In most cases, the identical product was found, but in some cases there were slight differences in model or presentation. In a small number of cases, no identical product was found, and a subjectively selected similar product was substituted.
3) The Keepa API was searched for the 500 products, using the product search query, which provides the first 10 results found on Amazon in response to a search for the product name. There were thus approximately 5,000 products found.
4) Information was obtained for these 5,000 products, including product characteristics (weight, size, color, etc.), historical prices, sales ranking, product categories and variations (in model, color, sizes, product grouping, etc.).
5) As most of these products have variations, and each variation has its own price history, the variations were also scraped for their price histories and particular characteristics. The final database included approximately 17,000 products (including the actual products and their variations), each with its own price history. It should be noted that the number of clothing products was quite large, as there were many combinations of models, sizes, and colors for each product.
6) The elimination of false positives from the search process (different products with similar names) greatly reduced the final number.

This procedure produced the following databases:

- A product database with information on physical characteristics, the date the product was listed on Amazon (which turned out not to be useful, as most products were listed at the time Keepa began to record the prices), the title and description of the product as listed by Amazon, as well as product features, color, model, size, brand, group, type, manufacturer, and ASIN (a unique Amazon product code).
- A price database with the date, price, and type of seller for each product ("New," sold by a third party and delivered by a third party or Amazon; or "Amz," sold and delivered by Amazon), and its ASIN.
- A variant database with information on the variants (different models, colors, sizes, etc.) of every product, which allows us to map the ASINs of products and their variants. The number of variants ranges from 0 to 648. If a product does not appear in this database, it has no variant. There are 14,933 products with at least one variant. Only one product has 648 variants.

Price Information from INEGI

Obtaining Publicly Available Data on Average Prices

The variable Consecutivo identifies a single product in a given month by Generic Type and Name of City. The variable Status helps to identify when the time series of a particular product is broken, and the price of a different product is recorded in its place. Specification, also a character field, helps to precisely identify the product. However, it is manually captured and therefore needs to be cleaned before use.

[33] Details of this selection are described in the next section.

We initially choose several classes of products. However, these are broad categories and may include more products than can be analyzed, so they are subsequently narrowed down. These are:

- Food, drinks, and tobacco
- Clothing, footwear, and accessories
- Furniture, appliances, and household accessories
- Health products
- Leisure products (entertainment items)

Data was downloaded[34] and cleaned.[35] INEGI follows a particular product across time. When one of these products changes presentation or is not found, the time series of that product is broken, so a new product is followed. This change in products over time also needs to be considered.

To determine the importance of the products to be analyzed, it is necessary to find an objective way to order them into a hierarchy. The weights used by INEGI to calculate the CPI are the best option; these are merged with the data.[36]

Selecting Products for Analysis

The next step is to obtain a database of time series of prices, at the city-product level: $price_{t,i,city}$. A deeper cleaning must be carried out to determine whether these products are the same across cities. The first wave of data cleaning is performed as follows:

- We keep only city-products for which we have at least 24 months of data.
- We keep only city-products with data recorded in 2017 and later, as we are interested in a change after that date.
- We remove products that are not sold on Amazon (e.g., frozen food, bulk food, perishable food, and oversized furniture).

Final Database

The INEGI information and Amazon information are merged on a product basis into the final database. By construction,[37] there can be several Amazon products for every INEGI product. These may be identical products, nearly identical products (for example, different models, colors, variations, packaging, or sizes), or very close substitutes (different brands of the same product category). We find that approximately 500 products from INEGI could be linked to approximately 5,000 corresponding products on Amazon. A single product from INEGI generally corresponds to several products Amazon. There are only five Amazon products that correspond to multiple products from INEGI.[38] In these few cases, the average price is taken of the INEGI products to maintain a 1-to-*n* relation (one INEGI product to *n* Amazon products).

---

[34] Data for 2018-21 is downloaded from https://www.inegi.org.mx/app/preciospromedio/?bs=18, and for 2011-18 from https://www.inegi.org.mx/app/preciospromedio/.

[35] There are problems with the description of some products in the csv files, which must be resolved with hard coding.

[36] Available at https://www.inegi.org.mx/programas/inpc/2018/. Some adjustments were needed to paste the data, as there are 32 generics with different names, but most are easily traceable.

[37] This is true because we manually picked approximately 600 products from the INEGI price database and searched for them (or for close substitutes) in Keepa.

[38] These few cases are mostly products which are nearly identical, but have minor variations in name or presentation, so one Amazon product corresponds to two or more from INEGI. For example, INEGI includes the prices of several varieties of Bimbo bread (under the Bimbo or Wonder brands, each with different types and presentations), but Amazon includes only the classic Bimbo bread.

This is a panel database, with a monthly time frame and a city-ASIN individual identifier.[39] The variables include the city corresponding to the INEGI price; the date of the observation (INEGI and Amazon); the ASIN Amazon assigns to the product; an identifier for the city of the INEGI product; the price of the INEGI product in the city that month (where there are two INEGI products with the same ASIN, we calculate the average, and if INEGI does not have information for that product in that city that month, the data is noted as missing); the monthly average price of the Amazon product (sold and delivered by Amazon or a third party). The database also includes the INEGI identifiers *generic*, *div*, *group*, *class*, *subclass*, *iddiv*, *idgroup*, *idclas*, and *idsubclass*, which represent different ways in which INEGI aggregates products and their identifiers; the weight of the INEGI product in the Mexican CPI; and the information and characteristics for each product from Amazon.

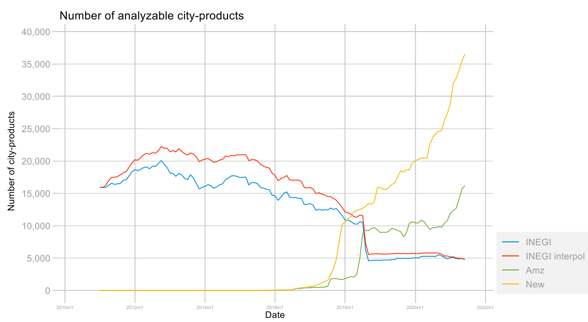
Brief Database Descriptive Statistics

There are some issues in the data that should be noted:

1) For several cities, information is only available after 2018.[40] However, most have data from January 2011. As would be expected, Mexico City includes the greatest number of products.
2) A large number of the matches found are in clothing categories.[41] Most of the clothing sold on Amazon includes multiple varieties, colors, and sizes, and every combination is considered a different product, usually with varying prices and availability. For this reason, the 500 INEGI products correspond to a much larger number of Amazon products: the number of matches in clothing tends to produce an overestimate.

Finally, Figures 9-11 show the number of analyzable products by month and by matches:

Figure 9. Number of Analyzable City-Products, All Categories, by Month.

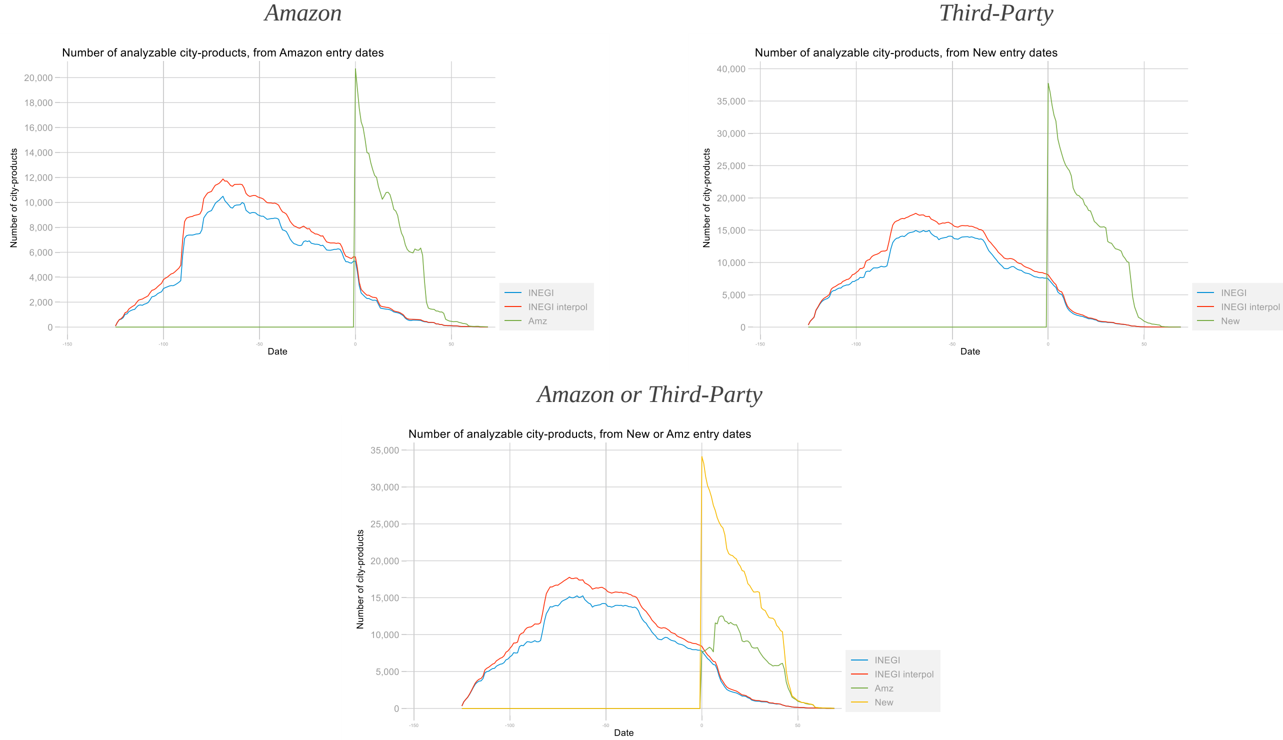


Note: Authors' calculations.

Figure 9 shows the number of analyzable city-products for all categories. Some INEGI series have important discontinuities, especially around 2018, when there was an important change in the products they followed. However, many can still be analyzed.

[39] As already noted, the 500 INEGI products are matched to 5,000 ASINs, in a perfect 1-to-*n* manner; each unique product has an ASIN. There are also different cities in the INEGI price data, but the Amazon price is the same for every city.
[40] These are Atlacomulco, Cancún, Coatzacoalcos, Esperanza (Sonora), Izúcar de Matamoros (Puebla), Pachuca, Saltillo, Tuxtla, and Zacatecas.
[41] These are mainly men's trousers and clothing, children's clothing, baby clothing, and shoes.

Figure 10. Number of Analyzable City-Products, by Type of Entry.



Notes: Authors' calculations, based on the microdata.

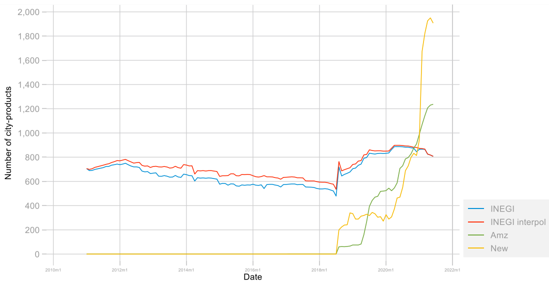
We note the following regarding Figure 10:

- As in Figure 9, there are important discontinuities around 2018, when there was an important change in the products INEGI followed.
- The Amazon or third-party series might cover too short a period, particularly in groceries, because Amazon had just recently started selling these types of products (often at the beginning of the pandemic, and in some cases only one year before the pandemic).

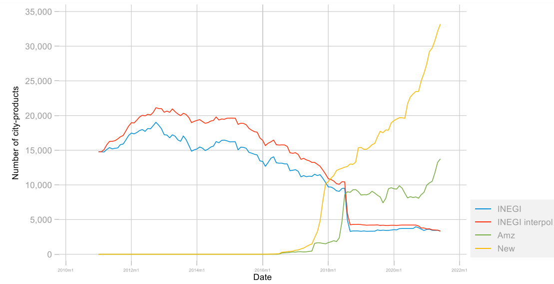
Figure 11 shows the results by product group. The number of products with available CPIs declines around 2018 because of the INEGI "base change" around August of that year. With this change, INEGI stopped following many of the products it had previously followed. The problem is particularly important in clothing, where many product series are lost.

Figure 11. Number of Analyzable City-Products, by Product Group

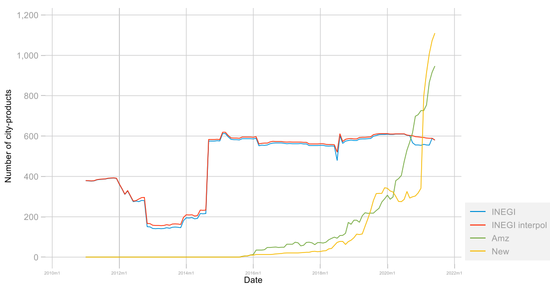
Food



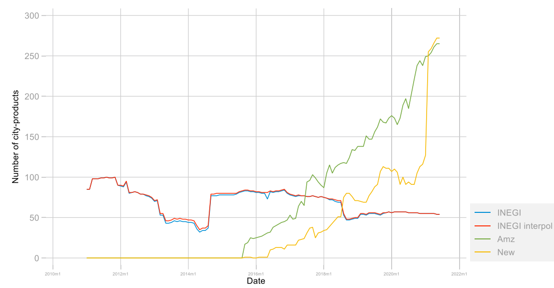
Clothing



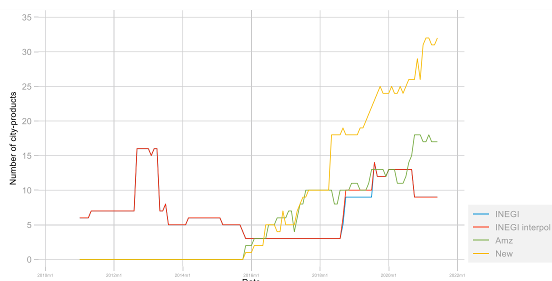
Furniture



Health



Leisure

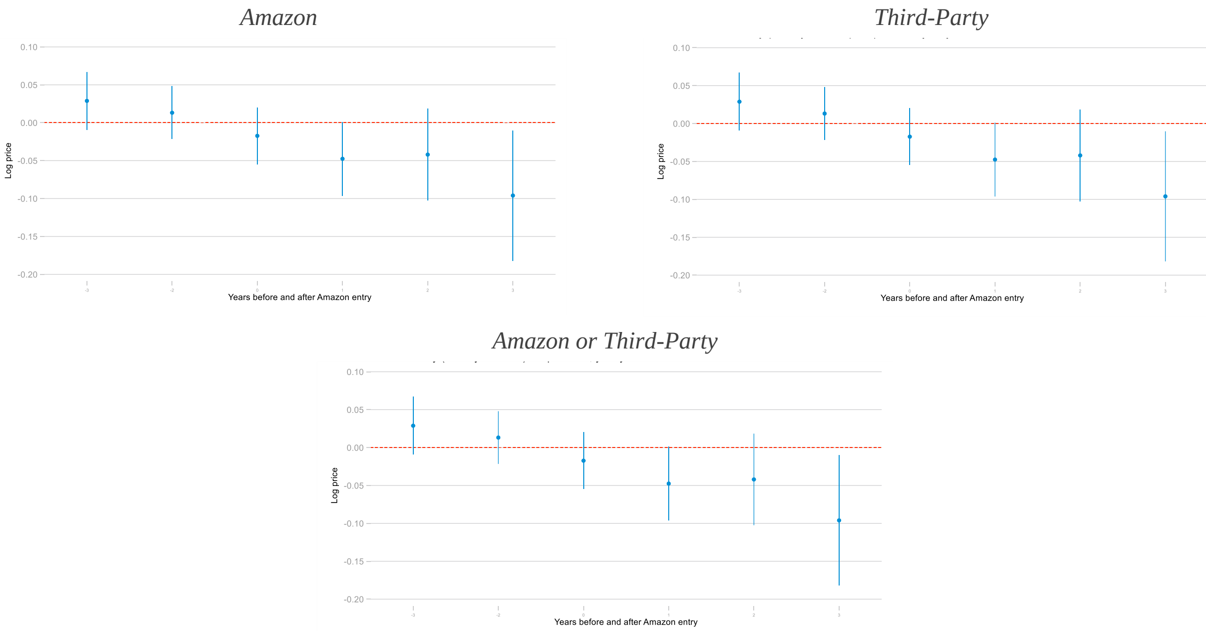


Note: Authors' calculations.

C. Additional Graphics, Yearly

Figure 12 shows yearly averages. The downward effect of prices, which is barely statistically significant, is around 2.5% a year for Amazon's entry; for third parties, it is not statistically significant. For Amazon or third parties, it is approximately 7% for four years, but not statistically significant.

Figure 12. Coefficients of the Fixed Effect Weighted Regression. Yearly. All products.

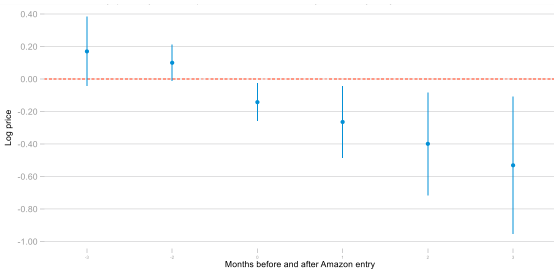


Notes: Authors' calculations, based on the microdata. This is the plot of the estimated coefficients: the event dummies, with 95% confidence intervals.

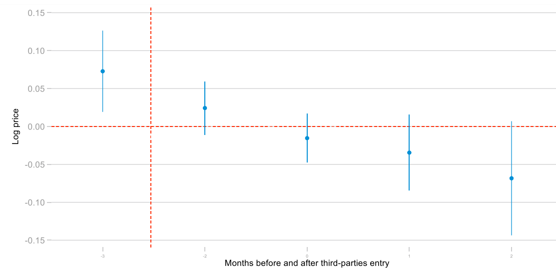
Figures 13-15 show the yearly results by product group. There is an important, very statistically significant downward effect, visible in Figure 13, for food, beverages, and tobacco, of approximately 15% a year. Third-party entry is not significant, but the effect is significant for Amazon or third parties and is approximately 10% a year. Figure 14 shows no statistically significant effect on furniture in any case. Figure 15 reveals a large and significant downward effect on the prices of health products, approximately 20% a year for Amazon entry, though it is not significant for third parties. There is a very large and significant effect of approximately 15% a year.

Figure 13. Coefficients of the Fixed Effect Weighted Regression. Yearly. 1. Food, Beverages, and Tobacco.

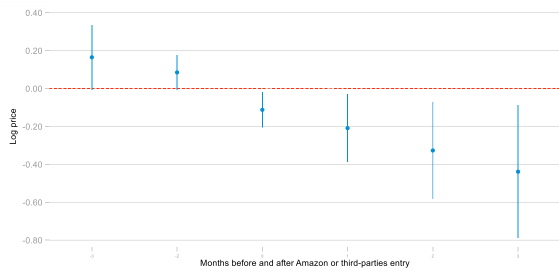
Amazon



Third-Party



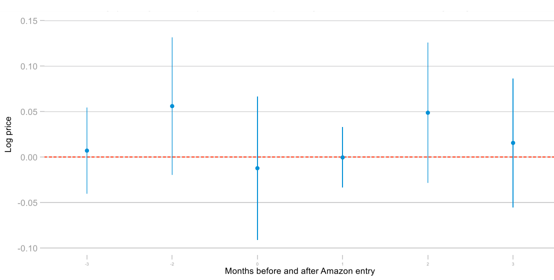
Amazon or Third-Party



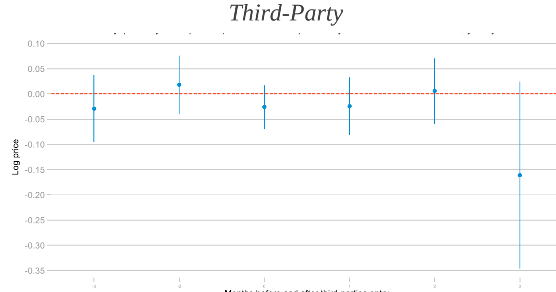
Notes: Authors' calculations, based on the microdata. This is the plot of the estimated coefficients: the event dummies, with 95% confidence intervals.

Figure 14. Coefficients of the Fixed Effect Weighted Regression. Yearly. 4. Furniture, Appliances, and Household Accessories.

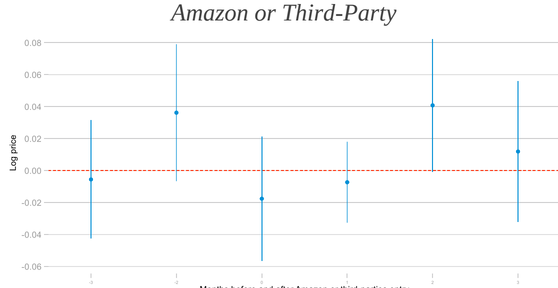
Amazon



Third-Party



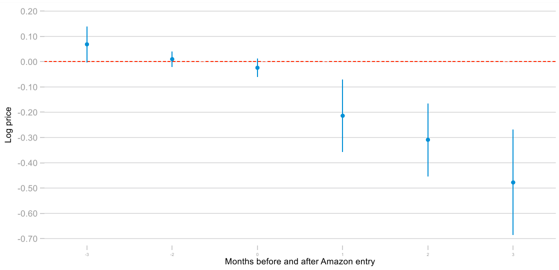
Amazon or Third-Party



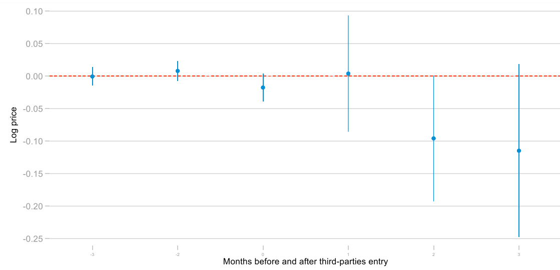
Notes: Authors' calculations, based on the microdata. This is the plot of the estimated coefficients: the event dummies, with 95% confidence intervals.

Figure 15. Coefficients of the Fixed Effect Weighted Regression. Yearly. 5. Health and Personal Care.

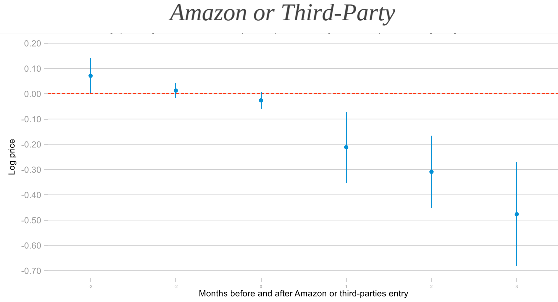
Amazon



Third-Party



Amazon or Third-Party



Notes: Authors' calculations, based on the microdata. This is the plot of the estimated coefficients: the event dummies, with 95% confidence intervals.

D. Regression Results of Event Analysis

Table 13 shows the regression for all products, monthly.

Table 13. Regression Results, All Products.

VARIABLE	**Amazon** log price	**Third Parties** log price	**Amazon or Third Parties** log price
-13	0.0581***	0.0453***	0.0324**
	(0.0144)	(0.0105)	(0.0128)
-12	0.0428**	0.0114	0.0226
	(0.0187)	(0.0155)	(0.0176)
-11	0.0341*	0.00406	0.0174
	(0.0193)	(0.0156)	(0.0179)
-10	0.0401**	0.00417	0.0191
	(0.0190)	(0.0156)	(0.0179)
-9	0.0389**	0.00247	0.0169
	(0.0183)	(0.0159)	(0.0180)
-8	0.0375**	0.0113	0.0245
	(0.0186)	(0.0144)	(0.0170)
-7	0.0354*	0.00970	0.0211
	(0.0185)	(0.0146)	(0.0173)
-6	0.0283	0.0107	0.0254
	(0.0186)	(0.0142)	(0.0167)
-5	0.0223	0.00726	0.0227
	(0.0177)	(0.0136)	(0.0159)
-4	0.0319*	0.00477	0.0259
	(0.0179)	(0.0135)	(0.0161)
-3	0.0127	0.00142	0.00615
	(0.0179)	(0.0136)	(0.0159)
-2	0.00796	0.00234	0.00105
	(0.0181)	(0.0144)	(0.0167)
0	-0.00166	-0.00128	-0.000238
	(0.0193)	(0.0129)	(0.0166)
1	0.00318	-0.00333	-0.000781
	(0.0183)	(0.0125)	(0.0157)
2	-0.00304	-0.00648	-0.00958
	(0.0187)	(0.0136)	(0.0163)
3	6.96e-05	-0.000846	-0.00272
	(0.0185)	(0.0149)	(0.0151)
4	0.00908	-0.00343	0.00276
	(0.0197)	(0.0146)	(0.0156)
5	-0.000254	-0.0129	-0.00213
	(0.0204)	(0.0122)	(0.0172)
6	-0.00562	-0.0166	-0.00629
	(0.0206)	(0.0131)	(0.0179)
7	-0.00821	-0.0278**	-0.0117
	(0.0200)	(0.0128)	(0.0174)
8	-0.0222	-0.0289**	-0.0199
	(0.0199)	(0.0142)	(0.0164)
9	-0.0228	-0.0268**	-0.0146
	(0.0215)	(0.0126)	(0.0178)
10	-0.0215	-0.0312**	-0.0140
	(0.0225)	(0.0140)	(0.0198)
11	-0.0168	-0.0328**	-0.0125
	(0.0265)	(0.0153)	(0.0229)
12	-0.0144	-0.0587***	-0.0205
	(0.0254)	(0.0145)	(0.0212)
13	-0.0530***	-0.0708***	-0.0472***
	(0.0172)	(0.0128)	(0.0141)
14	-0.0655***	-0.0672***	-0.0504***
	(0.0187)	(0.0143)	(0.0151)
15	-0.0827***	-0.0714***	-0.0695***
	(0.0177)	(0.0137)	(0.0149)
16	-0.0707***	-0.0591***	-0.0587***
	(0.0189)	(0.0167)	(0.0147)
17	-0.0732***	-0.0583***	-0.0554***
	(0.0180)	(0.0148)	(0.0138)
18	-0.0776***	-0.0827***	-0.0575***
	(0.0178)	(0.0171)	(0.0141)
19	-0.0736***	-0.0715***	-0.0576***

Table 13. Regression Results, All Products.

VARIABLE	**Amazon** log price	**Third Parties** log price	**Amazon or Third Parties** log price
	(0.0184)	(0.0154)	(0.0149)
20	-0.0688***	-0.0626***	-0.0482***
	(0.0202)	(0.0144)	(0.0152)
21	-0.0491***	-0.0363**	-0.0346**
	(0.0187)	(0.0155)	(0.0151)
22	-0.0504***	-0.0518***	-0.0406***
	(0.0177)	(0.0154)	(0.0146)
23	-0.0580***	-0.0622***	-0.0458***
	(0.0191)	(0.0162)	(0.0159)
24	-0.0814***	-0.0671***	-0.0515***
	(0.0190)	(0.0174)	(0.0148)
25	-0.0812***	-0.0400**	-0.0448***
	(0.0201)	(0.0159)	(0.0159)
26	-0.0683***	-0.0609***	-0.0447***
	(0.0179)	(0.0169)	(0.0141)
27	-0.0813***	-0.0532**	-0.0421**
	(0.0193)	(0.0258)	(0.0167)
28	-0.0667***	-0.0664***	-0.0380**
	(0.0226)	(0.0251)	(0.0161)
29	-0.0407*	-0.0807***	-0.0399**
	(0.0237)	(0.0229)	(0.0166)
30	-0.0698***	-0.0878***	-0.0621***
	(0.0225)	(0.0202)	(0.0163)
31	-0.0992***	-0.0742***	-0.0624***
	(0.0189)	(0.0202)	(0.0153)
32	-0.0986***	-0.0469**	-0.0575***
	(0.0210)	(0.0208)	(0.0169)
33	-0.0847***	-0.0602***	-0.0576***
	(0.0236)	(0.0197)	(0.0171)
34	-0.0440**	-0.0726***	-0.0412**
	(0.0224)	(0.0208)	(0.0172)
35	-0.0881***	-0.0792***	-0.0648***
	(0.0232)	(0.0172)	(0.0171)
36	-0.107***	-0.0332	-0.0570***
	(0.0218)	(0.0258)	(0.0199)
37	-0.0565***	-0.0542**	-0.0254
	(0.0207)	(0.0254)	(0.0162)
38	-0.0419	-0.0662*	-0.0257
	(0.0326)	(0.0364)	(0.0293)
39	-0.0710***	-0.0950***	-0.0432**
	(0.0252)	(0.0291)	(0.0220)
40	-0.0829***	-0.0776	-0.0421
	(0.0312)	(0.0598)	(0.0273)
41	-0.108***	-0.145***	-0.0735*
	(0.0397)	(0.0387)	(0.0377)
42	-0.0810*	-0.167***	-0.0478
	(0.0467)	(0.0301)	(0.0460)
43	-0.0531	-0.153***	-0.0206
	(0.0384)	(0.0388)	(0.0389)
44	-0.0826**	-0.209***	-0.0513*
	(0.0324)	(0.0550)	(0.0304)
45	-0.0908***	-0.238***	-0.0564**
	(0.0299)	(0.0467)	(0.0265)
46	-0.0802***	-0.151***	-0.0386
	(0.0279)	(0.0548)	(0.0251)
47	-0.145***	-0.292***	-0.108***
	(0.0276)	(0.0686)	(0.0246)
48	-0.140***	-0.122*	-0.103***
	(0.0320)	(0.0647)	(0.0285)
49	-0.129***	-0.216***	-0.0952***
	(0.0205)	(0.0349)	(0.0176)
Constant	3.405***	3.473***	3.483***
	(0.0140)	(0.0101)	(0.0124)
Observations	102,929	140,808	141,830
R-squared	0.994	0.994	0.994

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

Table 14 shows the regression for all products, yearly.

Table 14. Regression Results, All Products. Yearly.

VARIABLE	**Amazon** log price	**Third Parties** log price	**Amazon or Third Parties** log price
dummyamz1	0.0288	0.0466***	0.0209
	(0.0194)	(0.0169)	(0.0151)
dummyamz2	0.0132	0.0103	0.0102
	(0.0177)	(0.0136)	(0.0164)
dummyamz4	-0.0173	0.00124	-0.0126
	(0.0191)	(0.0116)	(0.0158)
dummyamz5	-0.0476*	-0.00441	-0.0310*
	(0.0249)	(0.0172)	(0.0169)
dummyamz6	-0.0420	-0.0171	-0.0278
	(0.0308)	(0.0300)	(0.0197)
dummyamz7	-0.0960**	-0.0301	-0.0745**
	(0.0438)	(0.0536)	(0.0318)
Constant	3.533***	3.592***	3.594***
	(0.0152)	(0.0354)	(0.0144)
Observations	10,222	13,963	14,064
R-squared	0.996	0.996	0.996

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

Table 15 shows the regression for all products, by product group, monthly.

Table 15. Regression Results, by Product Group.

	Amazon			**Third Parties**			**Amazon or Third Parties**		
VARIABLE	Food log price	Furniture log price	Health log price	Food log price	Furniture log price	Health log price	Food log price	Furniture log price	Health log price
-13	0.122***	0.00854	0.0356**	0.0582**	-0.000443	0.0487***	0.0310	-0.0112	0.0359**
	(0.0262)	(0.0130)	(0.0152)	(0.0254)	(0.00795)	(0.0140)	(0.0267)	(0.00762)	(0.0147)
-12	0.0894***	0.0444***	0.00993	0.0172	0.0100	0.0319*	8.16e-05	0.0268***	0.0157
	(0.0307)	(0.0146)	(0.0154)	(0.0357)	(0.00972)	(0.0172)	(0.0359)	(0.00935)	(0.0150)
-11	0.0747**	0.0428***	0.00429	0.0139	0.0123	0.0269	-0.00317	0.0257***	0.0112
	(0.0322)	(0.0149)	(0.0148)	(0.0353)	(0.00906)	(0.0187)	(0.0353)	(0.00906)	(0.0144)
-10	0.0710**	0.0448***	-0.00227	0.0153	0.0161*	0.00332	-0.00393	0.0270***	0.00253
	(0.0323)	(0.0147)	(0.0144)	(0.0355)	(0.00882)	(0.0168)	(0.0357)	(0.00911)	(0.0139)
-9	0.0723**	0.0399***	0.0106	0.0190	0.0150*	0.00741	0.00727	0.0216**	0.0105
	(0.0316)	(0.0132)	(0.0151)	(0.0337)	(0.00853)	(0.0200)	(0.0345)	(0.00842)	(0.0147)
-8	0.0602*	0.0357***	0.00985	0.0385	0.0138	0.0191	0.0257	0.0201**	0.00824
	(0.0333)	(0.0132)	(0.0167)	(0.0298)	(0.00936)	(0.0188)	(0.0323)	(0.00872)	(0.0163)
-7	0.0474	0.0327**	0.00461	0.0356	0.0115	0.0297*	0.0206	0.0176**	0.00306
	(0.0338)	(0.0136)	(0.0148)	(0.0293)	(0.00923)	(0.0174)	(0.0327)	(0.00871)	(0.0144)
-6	0.0475	0.0268*	-0.00644	0.0378	0.00739	0.0104	0.0227	0.0152*	-0.00749
	(0.0332)	(0.0139)	(0.0149)	(0.0301)	(0.00954)	(0.0168)	(0.0321)	(0.00868)	(0.0146)
-5	0.0352	0.0311**	-0.00732	0.0323	0.00612	0.0163	0.0177	0.0194**	-0.00600
	(0.0318)	(0.0133)	(0.0164)	(0.0279)	(0.0108)	(0.0182)	(0.0311)	(0.00847)	(0.0159)
-4	0.0332	0.0252*	0.00799	0.0153	0.00630	0.0228	0.0161	0.0134	0.00841
	(0.0314)	(0.0137)	(0.0149)	(0.0272)	(0.00960)	(0.0161)	(0.0305)	(0.00873)	(0.0146)
-3	0.0277	0.0169	-0.00387	0.00612	0.00263	0.00370	0.0100	0.00817	-0.00315
	(0.0318)	(0.0136)	(0.0142)	(0.0252)	(0.00930)	(0.0144)	(0.0304)	(0.00903)	(0.0139)
-2	0.0132	0.00650	0.00185	0.00149	-0.000470	0.0152	0.00677	-0.00286	0.00317
	(0.0316)	(0.0146)	(0.0157)	(0.0297)	(0.00893)	(0.0157)	(0.0316)	(0.00916)	(0.0155)
0	-0.0232	0.00712	-0.0205	-0.00942	0.00169	0.000593	-0.0110	0.00421	-0.0194
	(0.0313)	(0.0161)	(0.0176)	(0.0279)	(0.00971)	(0.0189)	(0.0309)	(0.00936)	(0.0173)
1	-0.0335	0.0110	-0.0103	-0.0187	0.00751	0.00217	-0.0238	0.00485	-0.0107
	(0.0291)	(0.0140)	(0.0182)	(0.0228)	(0.00915)	(0.0212)	(0.0298)	(0.00897)	(0.0178)
2	-0.0471	-0.000227	-0.00769	-0.0211	0.00705	-0.0148	-0.0310	-0.00372	-0.00940
	(0.0299)	(0.0137)	(0.0199)	(0.0241)	(0.00934)	(0.0227)	(0.0303)	(0.00926)	(0.0187)
3	-0.0534*	0.00497	-0.000968	-0.00146	0.000999	-0.0219	-0.0191	-0.00246	-0.00336
	(0.0305)	(0.0150)	(0.0236)	(0.0256)	(0.0101)	(0.0235)	(0.0307)	(0.00968)	(0.0204)
4	-0.0479	0.0138	-0.00143	0.000767	-0.00789	-0.000507	-0.0288	0.00467	-0.00393
	(0.0306)	(0.0150)	(0.0217)	(0.0263)	(0.0125)	(0.0239)	(0.0277)	(0.00971)	(0.0187)
5	-0.0483	0.00785	0.00316	-0.0369*	-0.00918	-0.0241	-0.0323	-0.00272	-0.000104
	(0.0311)	(0.0146)	(0.0285)	(0.0218)	(0.0149)	(0.0213)	(0.0280)	(0.0102)	(0.0222)
6	-0.0549*	0.00160	-0.0278	-0.0379*	-0.0158	-0.0304	-0.0345	-0.0117	-0.0249

Table 15. Regression Results, by Product Group.

	Amazon			**Third Parties**			**Amazon or Third Parties**		
	Food	Furniture	Health	Food	Furniture	Health	Food	Furniture	Health
VARIABLE	log price	log price	log price	log price	log price	log price	log price	log price	log price
	(0.0323)	(0.0146)	(0.0272)	(0.0225)	(0.0133)	(0.0288)	(0.0282)	(0.0104)	(0.0220)
7	-0.0725**	-0.00767	-0.00345	-0.0437*	-0.0163	-0.0488*	-0.0350	-0.0208**	-0.00742
	(0.0322)	(0.0148)	(0.0246)	(0.0229)	(0.0130)	(0.0274)	(0.0277)	(0.0103)	(0.0222)
8	-0.0900***	-0.0150	-0.0171	-0.0549**	-0.00299	-0.0320	-0.0432	-0.0237**	-0.0169
	(0.0328)	(0.0147)	(0.0189)	(0.0268)	(0.0136)	(0.0298)	(0.0302)	(0.0102)	(0.0174)
9	-0.109***	-0.0155	-0.0298	-0.0480**	-0.00236	-0.0588*	-0.0539*	-0.0156	-0.0288
	(0.0332)	(0.0162)	(0.0238)	(0.0232)	(0.0135)	(0.0340)	(0.0278)	(0.0116)	(0.0213)
10	-0.125***	-0.0177	-0.0352*	-0.0605**	-0.0134	-0.0728*	-0.0617**	-0.0217*	-0.0341*
	(0.0341)	(0.0175)	(0.0208)	(0.0238)	(0.0156)	(0.0438)	(0.0286)	(0.0128)	(0.0200)
11	-0.117***	-0.0271	-0.0176	-0.0503*	-0.0155	-0.0928*	-0.0522	-0.0356***	-0.0179
	(0.0379)	(0.0166)	(0.0192)	(0.0286)	(0.0129)	(0.0493)	(0.0329)	(0.0120)	(0.0182)
12	-0.0990***	-0.0252	-0.0242	-0.0714**	-0.0181	-0.111**	-0.0501	-0.0359***	-0.0256
	(0.0365)	(0.0179)	(0.0213)	(0.0313)	(0.0139)	(0.0463)	(0.0332)	(0.0122)	(0.0202)
13	-0.143***	-0.0299*	-0.0199	-0.0911***	-0.0277***	-0.127**	-0.0859***	-0.0325***	-0.0205
	(0.0329)	(0.0165)	(0.0205)	(0.0241)	(0.01000)	(0.0508)	(0.0272)	(0.0100)	(0.0191)
14	-0.153***	-0.0423**	-0.0453	-0.0511**	-0.0469***	-0.155***	-0.0773***	-0.0410***	-0.0436
	(0.0340)	(0.0182)	(0.0456)	(0.0237)	(0.0138)	(0.0455)	(0.0281)	(0.0117)	(0.0388)
15	-0.162***	-0.0544***	-0.137***	-0.0597**	-0.0354***	-0.195***	-0.0868***	-0.0569***	-0.137***
	(0.0342)	(0.0161)	(0.0285)	(0.0239)	(0.0124)	(0.0413)	(0.0278)	(0.0106)	(0.0279)
16	-0.149***	-0.0499**	-0.183***	-0.0733***	-0.0269*	-0.147***	-0.0832***	-0.0480***	-0.181***
	(0.0335)	(0.0195)	(0.0353)	(0.0249)	(0.0147)	(0.0508)	(0.0274)	(0.0106)	(0.0349)
17	-0.153***	-0.0539***	-0.129***	-0.0718***	-0.0257*	-0.0508	-0.0822***	-0.0385***	-0.129***
	(0.0336)	(0.0175)	(0.0492)	(0.0247)	(0.0155)	(0.0435)	(0.0270)	(0.0114)	(0.0484)
18	-0.162***	-0.0535***	-0.169***	-0.0657***	-0.0377**	-0.151**	-0.0829***	-0.0334***	-0.168***
	(0.0334)	(0.0147)	(0.0512)	(0.0250)	(0.0152)	(0.0754)	(0.0276)	(0.0101)	(0.0501)
19	-0.174***	-0.0496***	-0.107*	-0.0710***	-0.0397***	-0.168***	-0.0880***	-0.0440***	-0.109**
	(0.0328)	(0.0174)	(0.0570)	(0.0244)	(0.0140)	(0.0403)	(0.0270)	(0.0125)	(0.0553)
20	-0.184***	-0.0307	-0.164*	-0.0660***	-0.0199*	-0.200***	-0.0879***	-0.0265***	-0.164*
	(0.0333)	(0.0191)	(0.0972)	(0.0246)	(0.0111)	(0.0627)	(0.0287)	(0.00991)	(0.0961)
21	-0.191***	-0.0111	-0.133*	-0.0625**	0.00133	-0.0584	-0.0878***	-0.0157	-0.133*
	(0.0315)	(0.0189)	(0.0751)	(0.0251)	(0.0129)	(0.0548)	(0.0275)	(0.0105)	(0.0741)
22	-0.205***	-0.0191	-0.0881*	-0.0772***	-0.00806	-0.0597	-0.101***	-0.0178	-0.0880*
	(0.0349)	(0.0197)	(0.0512)	(0.0259)	(0.0101)	(0.0689)	(0.0281)	(0.0120)	(0.0502)
23	-0.219***	-0.0291	-0.0602	-0.0834***	-0.0242	-0.0879*	-0.112***	-0.0275*	-0.0610
	(0.0360)	(0.0202)	(0.0732)	(0.0252)	(0.0155)	(0.0529)	(0.0292)	(0.0145)	(0.0722)
24	-0.222***	-0.0389**	-0.125***	-0.0845***	-0.00614	-0.184*	-0.110***	-0.0196**	-0.126***
	(0.0388)	(0.0166)	(0.0429)	(0.0267)	(0.00936)	(0.102)	(0.0308)	(0.00967)	(0.0418)
25	-0.215***	-0.0317*	-0.137***	-0.0867***	0.0114	-0.118*	-0.101***	-0.00537	-0.137***
	(0.0401)	(0.0173)	(0.0419)	(0.0271)	(0.0115)	(0.0643)	(0.0318)	(0.0112)	(0.0411)
26	-0.202***	-0.0234	-0.142***	-0.107***	0.00232	-0.259***	-0.113***	-0.00216	-0.141***
	(0.0402)	(0.0151)	(0.0463)	(0.0276)	(0.0106)	(0.0999)	(0.0301)	(0.00985)	(0.0452)
27	-0.211***	-0.0329**	-0.271***	-0.118***	0.0254	-0.309*	-0.126***	-0.00290	-0.268***
	(0.0458)	(0.0162)	(0.0546)	(0.0301)	(0.0159)	(0.165)	(0.0316)	(0.0131)	(0.0527)
28	-0.224***	-0.00418	-0.345***	-0.123***	0.0237*	-0.375***	-0.133***	0.00865	-0.339***
	(0.0478)	(0.0228)	(0.0415)	(0.0305)	(0.0141)	(0.130)	(0.0315)	(0.0121)	(0.0393)
29	-0.265***	0.0420*	-0.345***	-0.128***	0.00881	-0.350***	-0.146***	0.0175	-0.337***
	(0.0453)	(0.0239)	(0.0414)	(0.0324)	(0.0135)	(0.0526)	(0.0326)	(0.0118)	(0.0414)
30	-0.253***	0.00935	-0.351***	-0.139***	0.00432	-0.268***	-0.147***	0.00174	-0.346***
	(0.0529)	(0.0225)	(0.0517)	(0.0343)	(0.0143)	(0.0867)	(0.0352)	(0.0114)	(0.0514)
31	-0.259***	-0.0437***	-0.285***	-0.150***	0.0162	-0.297***	-0.151***	-0.00637	-0.282***
	(0.0553)	(0.0157)	(0.0918)	(0.0311)	(0.0122)	(0.0910)	(0.0334)	(0.0108)	(0.0897)
32	-0.283***	-0.0341	-0.278***	-0.167***	0.0285*	-0.123	-0.165***	0.00138	-0.275***
	(0.0535)	(0.0211)	(0.0704)	(0.0318)	(0.0159)	(0.0928)	(0.0348)	(0.0136)	(0.0691)
33	-0.311***	0.0126	-0.401***	-0.158***	0.0185	-0.0611	-0.172***	0.0121	-0.399***
	(0.0549)	(0.0223)	(0.0654)	(0.0346)	(0.0118)	(0.114)	(0.0361)	(0.0107)	(0.0653)
34	-0.243***	0.0300	-0.309***	-0.157***	0.00846	-0.278***	-0.154***	0.0128	-0.308***
	(0.0664)	(0.0210)	(0.0468)	(0.0389)	(0.0145)	(0.0489)	(0.0411)	(0.0108)	(0.0464)
35	-0.268***	-0.00757	-0.374***		0.00282	-0.323***	-0.105**	3.43e-05	-0.373***
	(0.0555)	(0.0160)	(0.0756)		(0.00960)	(0.0703)	(0.0433)	(0.00874)	(0.0754)
36	-0.283***	-0.0112	-0.410***		0.0303**	-0.344***	-0.116***	0.0136	-0.409***
	(0.0543)	(0.0151)	(0.0543)		(0.0149)	(0.0726)	(0.0426)	(0.0103)	(0.0541)
37	-0.291***	0.0154	-0.316***		0.0243*	-0.298***	-0.119***	0.0202**	-0.316***
	(0.0547)	(0.0158)	(0.0510)		(0.0135)	(0.0528)	(0.0430)	(0.00952)	(0.0507)
38	-0.316***	0.0223	-0.302***		0.0144	-0.307***	-0.112*	0.0125	-0.302***
	(0.0538)	(0.0168)	(0.0754)		(0.0215)	(0.0584)	(0.0619)	(0.0114)	(0.0751)
39	-0.329***	-0.00971	-0.299***		0.00641	-0.380***	-0.124*	0.00375	-0.299***
	(0.0513)	(0.0184)	(0.0700)		(0.0118)	(0.0831)	(0.0635)	(0.0161)	(0.0699)
40	-0.346***	0.00121	-0.353***		0.0342**	-0.504***	-0.145**	0.0225	-0.352***
	(0.0568)	(0.0205)	(0.0745)		(0.0173)	(0.108)	(0.0602)	(0.0165)	(0.0742)
41	-0.360***	-0.00616	-0.441***		-0.0912**	-0.476***	-0.161**	0.00730	-0.440***

Table 15. Regression Results, by Product Group.

	Amazon			Third Parties			Amazon or Third Parties		
VARIABLE	Food log price	Furniture log price	Health log price	Food log price	Furniture log price	Health log price	Food log price	Furniture log price	Health log price
	(0.0528)	(0.0155)	(0.0890)		(0.0422)	(0.0816)	(0.0655)	(0.0127)	(0.0888)
42	-0.370***	-0.00602	-0.377***		-0.0429	-0.449***	-0.162**	0.00779	-0.376***
	(0.0539)	(0.0166)	(0.102)		(0.0533)	(0.0758)	(0.0672)	(0.0154)	(0.102)
43	-0.378***	0.0194	-0.308***		-0.0543	-0.495***	-0.163**	0.0327	-0.307***
	(0.0551)	(0.0200)	(0.0656)		(0.101)	(0.0563)	(0.0635)	(0.0210)	(0.0655)
44	-0.378***	0.0141	-0.345***		-0.0732	-0.507***	-0.210***	0.0118	-0.345***
	(0.0550)	(0.0244)	(0.0770)		(0.0567)	(0.0595)	(0.0503)	(0.0184)	(0.0768)
45	-0.382***	0.00404	-0.332***		0.0175	-0.523***	-0.217***	0.00476	-0.332***
	(0.0564)	(0.0233)	(0.0766)		(0.0251)	(0.0599)	(0.0497)	(0.0179)	(0.0762)
46	-0.391***	0.0134	-0.318***		-0.00814	-0.537***	-0.223***	0.0294*	-0.318***
	(0.0569)	(0.0206)	(0.0593)		(0.0256)	(0.0581)	(0.0499)	(0.0169)	(0.0589)
47	-0.434***	0.00770	-0.439***		0.0359	-0.533***	-0.245***	0.0204	-0.439***
	(0.0570)	(0.0197)	(0.0726)		(0.0276)	(0.0599)	(0.0493)	(0.0143)	(0.0720)
48	-0.445***	0.0186	-0.554***		0.0355	-0.548***	-0.249***	0.0233	-0.553***
	(0.0582)	(0.0178)	(0.0596)		(0.0274)	(0.0604)	(0.0550)	(0.0156)	(0.0591)
49	-0.424***	0.00519	-0.480***		-0.174***	-0.547***	-0.208***	0.0110	-0.480***
	(0.0558)	(0.0130)	(0.0594)		(0.0394)	(0.0606)	(0.0445)	(0.00887)	(0.0590)
Constant	3.044***	3.475***	4.184***	3.296***	3.481***	4.156***	3.325***	3.484***	4.184***
	(0.0245)	(0.0125)	(0.0140)	(0.0238)	(0.00759)	(0.0138)	(0.0251)	(0.00720)	(0.0136)
Observations	49,382	43,802	5,961	81,474	49,685	6,104	81,622	50,559	6,104
R-squared	0.990	0.999	0.992	0.992	0.998	0.991	0.992	0.999	0.992

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

Table 16 shows the regression for all products, by product group, yearly.

Table 16. Regression results, by Type of Product. Yearly.

	Amazon			Third Parties			Amazon or Third Parties		
VARIABLE	Food log price	Furniture log price	Health log price	Food log price	Furniture log price	Health log price	Food log price	Furniture log price	Health log price
-3	0.170	0.00720	0.0688*	0.0728***	-0.0293	-0.000547	0.165*	-0.00553	0.0713**
	(0.108)	(0.0242)	(0.0357)	(0.0273)	(0.0338)	(0.00715)	(0.0871)	(0.0189)	(0.0355)
-2	0.0998*	0.0561	0.0100	0.0243	0.0182	0.00799	0.0851*	0.0362*	0.0127
	(0.0566)	(0.0386)	(0.0156)	(0.0179)	(0.0294)	(0.00766)	(0.0467)	(0.0219)	(0.0154)
0	-0.143**	-0.0121	-0.0240	-0.0154	-0.0257	-0.0175	-0.112**	-0.0176	-0.0261
	(0.0590)	(0.0402)	(0.0183)	(0.0165)	(0.0218)	(0.0108)	(0.0474)	(0.0199)	(0.0164)
1	-0.265**	-0.000280	-0.214***	-0.0345	-0.0244	0.00391	-0.209**	-0.00728	-0.212***
	(0.113)	(0.0169)	(0.0728)	(0.0254)	(0.0291)	(0.0454)	(0.0908)	(0.0129)	(0.0713)
2	-0.399**	0.0489	-0.309***	-0.0683*	0.00596	-0.0958*	-0.327**	0.0408*	-0.309***
	(0.162)	(0.0393)	(0.0730)	(0.0384)	(0.0329)	(0.0491)	(0.130)	(0.0212)	(0.0726)
3	-0.531**	0.0157	-0.478***		-0.161*	-0.115*	-0.439**	0.0119	-0.477***
	(0.216)	(0.0361)	(0.106)		(0.0945)	(0.0675)	(0.179)	(0.0225)	(0.105)
Constant	3.025***	3.423***	4.231***	3.785***	4.062***	4.344***	3.367***	3.492***	4.229***
	(0.0804)	(0.0216)	(0.0246)	(0.0201)	(0.0240)	(0.00705)	(0.0642)	(0.0152)	(0.0236)
Observations	3,685	2,937	624	3,760	1,028	311	6,230	3,274	645
R-squared	0.991	0.999	0.994	0.998	0.999	0.998	0.993	0.999	0.994

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