Drip Pricing When Consumers Have Limited Foresight: Evidence from Driving School Fees*

Preliminary & Incomplete. Please do not circulate or quote.
Current version: 10/30/2012

David Muir†
Wharton School
University of Pennsylvania

Katja Seim‡
Wharton School
University of Pennsylvania

Maria Ana Vitorino§
Carlson School of Management
University of Minnesota

Abstract

This paper empirically investigates the add-on pricing behavior of firms in the Portuguese market for driving instruction. We develop a framework along the lines of Gabaix and Laibson (2006) where consumers purchase a base and, with some probability, an add-on product from the same firm, but are not always aware of the possible need for the add-on product. We show that a typical loss-leader pricing strategy emerges in horizontally differentiated markets whereby markups on the base product are artificially lowered, while firms price the add-on at monopoly levels. We test the implications of the model using a detailed snapshot of industry data on student characteristics and preferences, school attributes including prices and costs, and market demographics for a cross-section of local markets with differing numbers of school competitors. We find significant evidence in support of our model predictions, including that firms face a substantial profit motive in the add-on market. Most notably, markups for the base product, but not the add-on products, decline in the number of competitors a firm faces, a prediction that has not been established in the literature to date.

Keywords: market structure, loss-leader pricing, myopic consumers

JEL Classification: L10, L15, L80

*We thank Susana Paulino at Instituto da Mobilidade e dos Transportes Terrestres (IMTT) for access to the data and for her generously provided time in giving insights and answering questions about the industry. Ana Isabel Horta provided excellent research assistance and Susana Belo assisted with local data collection. We gratefully acknowledge support from the Dean's Research Fund and the Global Initiatives Research Program at the Wharton School. All errors are our own.

†Department of Business Economics & Public Policy, Philadelphia, PA 19104, muir@wharton.upenn.edu.
‡Department of Business Economics & Public Policy, Philadelphia, PA 19104, kseim@wharton.upenn.edu.
§Marketing Department, Minneapolis, MN 55455, vitorino@umn.edu.

1 Introduction

In many industries, sellers advertise the low price for a base or upfront product in hopes of generating subsequent sales of other "add-on" products in greater numbers or at greater profit margins than the upfront product's sales. Rational-actor models typically explain the notion that multi-product firms might sell some products as low-markup "bargains" or loss leaders, only to recover these losses with high-markup "ripoffs" (Holton 1957) by search costs to learn unadvertised add-on prices, switching costs after the upfront product's purchase, or price discrimination in selling vertically differentiated upfront and add-on products (see Ellison 2005 for a general framework that intersects these three explanations).

Gabaix and Laibson (2006) formalize another explanation for high mark-ups in add-on markets: shrouding of high add-on prices may be optimal when consumers are naïve and do not anticipate that the price of the unadvertised add-on product is likely to be high. As examples, they point to tied products such as printers and toner cartridges where consumer unawareness of the high markup of cartridges reduces the role that the tied product plays in the purchase process.

The goal of this article is to investigate firm competition in the upfront and add-on markets when consumers, instead of being unaware of the high prices of add-ons, systematically underestimate their demand for the add-on. Such limited foresight is likely of importance in consumer financial markets, where consumers might underestimate their future usage of additional features of bank or credit card accounts, such as overdraft services or financing, at the time when they open an account and focus on monthly fees only in selecting between competing offers (Stango and Zinman 2009) or in professional testing and certification markets where people might systematically overestimate their likelihood of obtaining a certification at the first try or might not be aware of the frequency of re-certification requirements. Limited foresight can be important in a business-to-business context as well. Upstream manufacturers might sell new capital to downstream wholesalers at below-cost prices, only to recoup upfront losses from their myopic customers – often those who have not had repeated, contracted interactions with the upstream vendor and who underestimate the unavoidable future maintenance and replacement part expenses in add-on markets.

With this goal in mind, we set up a simple model that builds on Gabaix and Laibson (2006) model of add-on pricing with myopic consumers. On the demand side, we allow for two types of consumers who purchase an upfront and, with some probability, an add-on product from the same firm. Sophisticated consumers are aware of the possibility of needing to purchase the add-on, and are able to engage in costly effort to reduce their purchase incidence. Myopic consumers, on the other hand, believe that their demand for the add-on

product is nonexistent and do not take its price into account when choosing their supplier. We show that a typical loss-leader pricing strategy emerges whereby the upfront product is sold at below the usual profit margins, while firms price the add-on at supra competitive levels.

As in Gabaix and Laibson, the assumption of Bertrand competition among symmetric firms results in profit neutrality across the two products, with add-on profits offsetting losses in the upfront product market. This has two consequences for the upfront product's price. First, a larger share of myopic consumers, who unexpectedly (to them) participate in the add-on market and generate profit there, depresses the upfront product's price. Second, increases in the probability with which consumers require the add-on product similarly raises profitability, which is offset by a lower upfront product price and markup.

The majority of the paper empirically investigates firm pricing behavior in a typical add-on market. We study competition in the Portuguese market for driving instruction. Driving schools offer both courses of instruction to prepare students for the initial driving exam – our base or upfront product – and make-up courses for students who fail either the theory or the practice exam at the first go – the add-on product. Exam repetition is common; 45.2 percent of students fail either the theory exam or the practice exam the first time they take it. At the same time, the Portuguese regulator, the Instituto da Mobilidade e dos Transportes Terrestres (IMTT), asserts that a significant share of students is unaware of the high fail rates when they enroll in a school. Since students do not engage in repeat purchases, attending driving school only once, there is limited scope for uninformed students to understand schools' pricing practices and their relevance to them.

The high fail rates result in the add-on market being a significant source of revenue and profit to schools; our data indicate that for the average school, 13.9 percent of school revenue and an estimated 23 percent of variable profit derives from exam repeat courses. In monetary terms, schools earn an additional €0.66 for every €1 the average student generates in profit from its upfront service.

A general challenge in empirically studying add-on pricing and price shrouding lies in the fact that sufficiently rich data, where we observe the entire sequence of purchases consumers make and the prices they pay for each purchase, are difficult to obtain. A unique feature of our data is that we observe the universe of Portuguese driving students over a three-year period, including information on the driving school they attend, the sequence of their driving exam outcomes, and demographic information such as their age, gender, and place of residence. We combine these data with information on school characteristics and hand-collected prices schools charge for the basic driving course, as well as for repeating either the theory or the practice exam, which legally requires that students complete a fixed number of

additional lessons. These detailed data allow us to estimate the marginal cost to the school of each student who enrolls in the basic driving course, the repeat course for the theory exam, and the repeat course for the practice exam, and to compute profit margins for the upfront and add-on products. We observe individual demand and firm behavior in a cross-section of local markets with differing numbers of school competitors, which allows us to study add-on pricing in markets with varying competitive environments.

Our setting exhibits a number of other features that renders it amenable to studying add-on pricing when consumers are imperfectly informed about the add-on. First, since repeat courses consist of additional lessons only, and thus are effectively fractional versions of the upfront product, the products exhibit the same cost structure. Efficiency reasons can be largely ruled out as drivers for differential pricing of the two products. Second, the vast majority of – if not all – students would prefer not to purchase the product, but are forced to based on an arguably objective decision rule. Thus, we do not need to disentangle whether consumers buy the add-on purely because they were not fully informed about the product at the time of the initial purchase commitment or because they value it highly and would have bought it regardless.

Descriptive evidence from our setting and its rich data corroborates a number of our model predictions. While the average Lerner index in the upfront market averages 13.6 percent, the corresponding indices in the theory and practice add-on markets average 85.8 and 47.7 percent, respectively. That is, while schools earn low markups in the upfront market, they simultaneously reap supra competitive markups in the add-on markets. Moreover, variation in the upfront markups across schools – but not in the add-on markets – stems from the variation that different market structures impose on the competitive environments in different municipalities. Both OLS and two-stage least squares regressions of each school's markups on its municipality's market structure, as well as a number of school- and municipality-related controls, lend proof to such predictions.

Schools' upfront market markups strongly correlate with the number of schools in the municipality in all model specifications, an outcome that alludes to the standard Bertrand result in which competition in prices propels prices downward toward competitive levels. Specifically, for every school added to a school's municipality, upfront markups decrease by about €9, or 7.8 percent, on average. On the other hand, schools' add-on markups remain remarkably uncorrelated with the market structure across all specifications, a facet of competition, or the lack thereof, predicted by our theoretical model. Schools benefit from highly localized market power over locked-in students in retake markets and feel scant effects of competitive pressures from other schools in the municipality.

One additional facet of our model that deserves attention is the profit neutrality outcome

that is standard in loss-leader pricing models: firms collect the same aggregate profit, but shift how that profit is distributed over the add-on and upfront markets. With myopic students, however, there are still important distributional consequences despite the profit neutrality of schools. If sophisticates are able to reduce their demand in the add-on market, they reap the benefits of low prices in the upfront market. Myopes, on the other hand, who have limited foresight about their potential future need for the add-ons, more frequently pay add-on surcharges in equilibrium. Thus, myopic students - whose business in both markets comprises a significant portion of schools' profits - effectively subsidize sophisticates with lower prices in the upfront market.

The mechanism behind such findings blends ideas of several literatures. Our theoretical result that schools have an incentive to charge supra competitive prices in the add-on market – and that such high markups persist in markets with numerous competitors – harkens back to Diamond (1971) who first proposed that the Bertrand paradox is broken as soon as search costs are introduced into an otherwise homogeneous market. In models with either spatial (Lal and Matutes 1994) or independent spatial and vertical differentiation between firms (Verboven 1999), search costs entail that consumers find it too exogenously costly to observe unadvertised add-on prices before they purchase the advertised base good. While in these models, any positive profits earned in the aftermarket are competed away by losses incurred from the loss leader market, Ellison (2005) shows that strictly positive profits can be earned in equilibrium if vertical tastes are correlated with the strength of horizontal preferences. When consumers cannot observe the price of the higher quality add-on, obfuscation then also serves to facilitate second-degree price discrimination (Stole 2007). Our model generates high markups for the add-on product without the presence of such exogenous search costs.

Instead, as in classic repeat purchase models of pricing with switching costs (Klemperer 1987a,b; Beggs and Klemperer 1992), we assume that consumers are locked into purchasing both the upfront course and a possible repeat course from the same driving school. An inability on the part of firms to commit to prices for subsequently purchased add-on products in the initial period is one possible explanation for loss-leader pricing: when firms cannot commit to prices – and consumers thus cannot observe them when they purchase the upfront product – they exercise ex-post monopoly power and set add-on prices above marginal cost once a consumer is locked in (Farrell and Klemperer 2007). In our setting, firms are required to keep at hand a full schedule of prices. Our model thus posits that students have infinite switching costs, but we otherwise assume that schools commit to the add-on price when setting the price menu in the initial stage of the game. The typical loss-leader strategy arises instead because consumers only partially account for the add-on price during the school choice process; as a result schools compete primarily over the upfront price.

Gabaix and Laibson (2006) form part of a broader emerging literature that explores how biases in consumers' beliefs explain various pricing practices. Some models explain subscription pricing with quasi-hyperbolic discounting (Oster and Scott Morton 2005; DellaVigna and Malmendier 2006), use loss aversion to motivate bait-and-switch pricing (Köszegi and Rabin 2006; Heidhues and Köszegi 2010), apply boundedly-rational heuristics to study how market equilibria exploit noise in consumer product evaluations (Jin and Leslie 2003; Spiegler 2006), and demonstrate the impact on demand of alternative price frames (Wertenbroch 1998; Bertrand et al. 2010).

We also contribute to a burgeoning literature on price shrouding. Early work (Milgrom 1981; Jovanovic 1982) illustrates that, in the absence of frictions such as switching and search costs, homogeneous firms have an incentive to disclose all information relevant to the consumer's choice process unless consumers are boundedly rational. Then, add-on prices might be shrouded in equilibrium if consumers are myopic (Gabaix and Laibson 2006[1]), lack self-control (DellaVigna and Malmendier 2004), are overconfident (Grubb 2009), or vary in their tastes for add-ons and advertising add-on prices is expensive (Ellison 2005). With myopic consumers, Spiegler (2006) further demonstrates that firms may not have an incentive to educate consumers about their competitors' shrouded prices. While driving schools do not explicitly shroud their prices, that most students do not seek the full price menu at the time of purchase of the upfront service alludes to the implicit shrouding behavior of firms over consumers who lack perfect foresight of future contingencies.

The empirical literature of loss-leader and specifically shrouding models is scant. Walters (1988) establishes a surprisingly limited impact that loss leaders have on grocery store traffic. More recently, several papers' findings are consistent with products serving as loss leaders (Chevalier et al. 2003; Ellison and Ellison 2009). In studying online computer memory chip purchases, Ellison and Ellison (2009) show that a loss-leader firm that shrouds add-ons is profitable by attracting a large number of customers who end up buying upgraded memory chip products at higher prices. In a field experiment, Chetty et al. (2009) find that demand falls when retailers post tax-inclusive prices for personal care products and find a similar response in historic state-level alcohol consumption when retailers disclose taxes in alcohol prices. Hossain and Morgan (2006) and Brown et al. (2010) demonstrate that raising shrouded shipping charges increases revenues and the number of bidders attracted to eBay auctions in a field experiment.[2] The UK's Office of Fair Trading 2010 conducted a market study to investigate the phenomenon of add-on or "drip" pricing, and presented

[1] A number of authors, including Miao (2010), Kosfeld and Schüwer (2011), and Heidhues and Murooka have confirmed the results in Gabaix and Laibson (2006) in more general settings.

[2] This contrasts with mixed results from laboratory experiments (Morwitz et al. 1998; Bertini and Wathieu 2008).

evidence that such price partitioning leads consumers to make choices they would not have otherwise made if the prices had not been intentionally manipulated. Moreover, their study found that, of the six potentially manipulative pricing strategies that firms can adopt, drip pricing generates the greatest potential harm for consumers. Our paper adds to this small but growing empirical literature, but is unique in that it uses field data to examine the role of competition in affecting add-on pricing.

The paper proceeds as follows. Section 2 presents a basic loss-leader model with boundedly-rational students. Section 3 introduces the data and the institutional setting of Portuguese driving schools. Section4 provides evidence to support some of the basic predictions and implications of the model. Section 5 concludes.

2 A Model of Add-on Pricing with Myopic Students

Here, we outline a simple model of add-on pricing in the spirit of Spiegler (2011) and Gabaix and Laibson (2006) to illustrate that the loss-leader pricing strategies common to multi-product settings with search or switching costs can also arise when consumers have limited foresight of their demand for the add-on product.

Consider a two-period model of a market with n symmetric schools and a continuum of students. The schools offer an upfront or base service u – a course of instruction to prepare for the driving exam – and an add-on service a for an additional fee – a make-up course for exam re-takers. We assume that the add-on price takes the form of a surcharge: students who fail the exam do not have a choice but to purchase the add-on from their chosen school.[3]

In period one, each school j simultaneously chooses and commits to a pricing strategy $(p_{u,j}, p_{a,j})$, where $p_{u,j}$ and $p_{a,j}$ are the prices of school j's upfront and add-on services. As in Gabaix and Laibson (2006), the add-on price p_a is effectively bounded above by $\overline{p}_a$. For example, if a student is forced to pay a high repeat-course price, the student might choose not to continue with driving instruction or lodge a complaint with IMTT. While IMTT does not directly regulate the price of repeat courses, its oversight likely also limits the fees that schools can charge. Firms face constant and nonnegative marginal costs of providing each service, (c_u, c_a). Students then choose a school to enroll in and purchase the upfront service u. In the second period, students learn whether or not they need to buy the add-on at its set price depending on their exam results. Students have a strictly positive, identical probability of $\overline{\lambda} \in (0, 1]$ of failing the initial driving exam.

We assume that there are two student types in the market: a share of $\alpha \in (0, 1)$ myopes m

[3]We discuss below tha,t in our data, we observe a negligible share of school transfer by students, equivalent to the lock-in assumed here.

and $(1-\alpha)$ sophisticated students s. In period one, myopes have limited foresight about the add-on service; they only become aware of it ex-post when they fail the exam and are forced to purchase it. The assumption of limited foresight might reflect that myopic students assign zero probability to failing the exam – the students suffer from over-optimism, for example – or that the students are unaware of the existence of makeup course fees simply because the school does not prominently advertise this information.

In contrast, sophisticated students recognize the possibility of having to retake the exam. We assume that they form rational expectations over their probability of failing the exam, assessing it correctly at λ, and consider the add-on service when making their choice in the first period. Sophisticates can engage in costly effort to reduce their probability of failing from $\overline{\lambda}$ to $\underline{\lambda} > 0$ at an effort cost e.[4]

In line with our empirical setting, we assume, as in Gabaix and Laibson (2006), that students make a discrete school choice, allowing for heterogeneous valuations of each school. For simplicity, we assume that there are no systematic differences in valuations by type. The utility of student i of type $\{m, s\}$ from enrolling at school j is given by

$$\begin{aligned} u_{ij}^{m} &= v - p_{u,j} + \epsilon_{ij} \\ u_{ij}^{s} &= v - p_{u,j} - \lambda p_{a,j} + \epsilon_{ij}, \end{aligned} \tag{1}$$

where ϵ denotes the student's heterogeneous valuation for the school, such as the distance he travels to the school, and $\lambda = \{\underline{\lambda}, \overline{\lambda}\}$, depending on whether the sophisticate chooses to engage in effort.[5]

A student's demand for each school's services is given by the probability that the (expected) utility of school j exceeds that of all competing school $k \neq j$. Under the assumption that ϵ is distributed extreme value, this results in common multinomial Logit school choice

[4]For example, sophisticates might study more for the theory exams or try harder at their practice lessons so as to minimize their probability of failing and the likelihood that they will be required to purchase the add-on service. Unlike similar models, whereby sophisticates face a perfect substitute for the add-on service, however, we assume that sophisticates cannot reduce their probability of needing to purchase the add-on service to zero despite their best efforts to avoid it.

[5]Note that we do not model the school's choice of whether to unshroud or publicize the upfront prices. We show in the Appendix that, in our setting, where add-ons are unavoidable, the shrouded and unshrouded equilibria are equivalent if sophisticates form rational expectations about add-on prices and unshrouding is costless. We therefore simply assume that sophisticates are aware of the add-on price and form expectations solely with regard to their demand for the add-on, but not the firms' prices.

probabilities:

$$\begin{aligned} D_j^m &= \left(1 + (n-1)\exp\left(\frac{p_{u,j} - p_{u,-j}}{\sigma}\right)\right)^{-1} \\ D_j^s &= \left(1 + (n-1)\exp\left(\frac{p_{u,j} - p_{u,-j} + \lambda(p_{a,j} - p_{a,-j})}{\sigma}\right)\right)^{-1}, \end{aligned} \tag{2}$$

denoting as σ the scale parameter of the extreme value distribution.

Consider first pricing in the add-on market. Since students are locked in to their school upon failing the initial exam, the school serves as a monopolist over its demand and optimally charges the highest possible price, $\overline{p}_a$.

Now consider pricing of the base service. We analyze the symmetric case, and assume that firms compete on price. In the Appendix, we prove that there is a unique symmetric equilibrium characterized by the following pricing strategies:

$$\begin{aligned} p_{a,j}^* &= \overline{p}_a \\ p_{u,j}^* &= c_u + \frac{\sigma n}{n-1} - \left((1-\alpha)\overline{\lambda} + \alpha\underline{\lambda}\right)(\overline{p}_a - c_a) \end{aligned} \tag{3}$$

provided effort costs e are at most equal to $(\overline{\lambda} - \underline{\lambda})\overline{p}_a$. We make the following observations about the equilibrium prices:

1. Schools set the price of the add-on service to be the same and equal to the ex-post monopoly price for the add-on service.

2. The add-on price does not depend on the fraction of myopes or probability of failing because demand is perfectly inelastic in the aftermarket.

3. The price of the upfront service is non-monotonically decreasing in the fraction of myopes, or

$$\frac{\partial p_u^*}{\partial \alpha} = \left(\underline{\lambda} - \overline{\lambda}\right)(\overline{p}_a - c_a) \leq 0 \tag{4}$$

   Thus, as the fraction of myopes increases, the upfront price decreases to reflect that schools anticipate larger profits from the add-on service; schools want to attract more myopes upfront with their loss leader service, only to recoup these losses in the aftermarket.[6]

[6]Note that this conclusion depends on the relative ordering of the probabilities of failing by myopes and sophisticates. If sophisticated students had a strictly higher probability of failing than myopes regardless of effort expended, the price of the upfront service would increase in the share of myopes.

4. The price of the upfront service is monotonically decreasing in the type-weighted probability of failing the exam, $\left((1-\alpha)\overline{\lambda}+\alpha\underline{\lambda}\right)$. The greater the numbers of students schools can attract in the aftermarket, the lower the price they charge in the upfront market to entice them to choose its service in the first place.

5. Finally, markups in the upfront market decline as the number of firms increases. With large n and a sizable aftermarket, the model allows for the possibility that margins in the upfront market are negative and are offset by large positive markups in the add-on market.

The equilibrium pricing strategies in Equation 3 reflect that sophisticated students engage in costly effort to reduce their exposure to the add-on market. Such effort is not necessarily efficient. The firm foregoes profit in the amount of $(\overline{p}_a - c_a)\left(\overline{\lambda}-\underline{\lambda}\right)$ on every sophisticate. The choice to engage in effort is thus only efficient if $e \leq c_a\left(\overline{\lambda}-\underline{\lambda}\right)$, whereas the student's choice to do so reflects the prices the student pays for, rather than the cost of providing, the add-on service.

In section 4 we investigate the extent to which these properties of the equilibrium pricing strategies are borne out by the pricing patterns we observe in our data. Beforehand, we briefly describe driving instruction and schools in Portugal and summarize our sources of data.

3 Background and Data

3.1 The Portuguese Market for Driving Instruction

We begin with an overview of the process of obtaining a driver's license in Portugal, the market for driving instruction, and the role of the IMTT as the regulatory agency that oversees driving instruction.

To obtain a driver's license, learner drivers aged 18 years or older must first enroll in an IMTT-authorized driving school.[7] There, they must complete 28 theory lessons, the curriculum of which is set by IMTT, and a minimum of 32 on-road driving lessons. After completing the required theory lessons, the students take a computerized theory exam. Subsequently, they perform an on-road driving test. Both exams are administered at one of 35 exam centers. Twenty-two exam centers are managed by IMTT, while private organizations operate the remainder. An IMTT certified examiner oversees the on-road driving test; as of 2011,

[7]Initial authorization by IMTT requires, among others, proof that the proposed owner has at least five years of experience in driving instruction, that the school is financially viable, and that the fleet and facilities satisfy certain IMTT standards.

there are 573 active examiners.[8] Upon successful completion of both exams, the IMTT issues a probationary license to the new driver. The license converts to an unrestricted license automatically should the driver avoid serious driving violations during an initial five-year period. In 2011, IMTT issued a total of 95,556 new driver's licenses, approximately 85% of which are category B passenger vehicle licenses on which we focus. For these students, the process from registering at a school to receiving the probationary license takes 8.8 months to complete on average.

As of the end of 2011, there are 1,119 driving schools on mainland Portugal and 27 schools in the archipelagos of Madeira and the Azores. The industry's current size reflects significant growth in the number of firms since 1998, when only 527 schools served the population of 9.9 million. Entry was the result of significant liberalization efforts of professional services in general in the late 1990s. Previously, the regulatory framework restricted the number of schools serving each municipality to be a step function of the municipality's population and placed ceilings on school fees for instruction. Legislative reforms lifted both price and entry restrictions in 1998, resulting in an immediate increase in the number of schools by 25 percent and a stabilization at current levels by 2005, with virtually no exit of existing schools. A number of regulatory restraints remain in place, including territorial restrictions limiting all business to be conducted within the municipality covered by the owner's license and regulations governing the sharing of resources between commonly-owned schools that limit the presence of multi-outlet chains of schools. 87% of owners operate a single school only, and another 9% operate two schools. Figure 1 plots the locations of all driving schools in mainland Portugal by municipality population density.

==========

Insert Figure 1 about here

==========

Since the deregulation in 1998, each school sets its own price for services. A flat fee covers the base course of instruction, including classroom time, materials, practice theory exams, on-road driving lessons, and any fees that the student has to pay to the exam center to take the two exams. The schools handle the submission of fees to the exam center, as well as the scheduling of exams. If the learner driver fails the theory exam, he must pay a fee to retake the exam. Similarly, should a student fail the practice exam, he must pay an extra fee to complete an additional five driving lessons and to re-take the practice exam. Applicants are

[8]The majority of examiners work at a single exam center only. In an effort at reducing possible examiner corruption, the IMTT periodically rotates a share of examiners across its own 13 exam centers. In aggregate, however, only nine percent of active examiners are rotating examiners as of 2009.

entitled to a total of five attempts at the practice exam over the course of two years; no limits are placed on the number of attempts at passing the theory exam. We summarize the schools' fees for the base course, retaking the theory exam, and preparing for and retaking the on-road driving test, as well as the frequent instances of failing either of the exams next.

3.2 Data

Our empirical analysis combines a number of data sets at the school, student, and municipality-level. We obtained data on non-price attributes of schools from the IMTT. These data, which are collected as part of each school's licensing process and which are updated periodically, include information about the school's facilities, the number and tenure of its instructors, the number of license classes offered, and characteristics of its fleet of instruction vehicles.[9] We geocoded school addresses using GIS software to derive detailed location information. Lastly, we added hand-collected school prices and estimated costs. More details about our price and cost data follow below.

Second, we obtained individual student-level school enrollment and driving exam information from the IMTT. The comprehensive database includes the full universe of learner driver candidates who applied for driving instruction between 2008 and 2010. We observe the date on which the student applies for his learner's permit and the permit's license category, the dates, times, exam centers and outcomes of each exam, the final licensing date, as well as the candidate's age, gender, and home zip code. We assume that each student resides at the centroid of his zip code area to compute his distance to his chosen school.

Finally, we obtain demographic and other data at the level of the municipality from the market research company, GrupoMarktest, the Ministério do Trabalho e da Solidariedade Social, and Statistics Portugal.[10]

3.2.1 Price Collection and Sample Markets

While the IMTT collects various pieces of information about the schools it oversees, it does not collect price or fee information. We thus complement their data with hand-collected, detailed price information. To keep the data collection process manageable, we focus on 13 of the 18 Portuguese districts, excluding most notably the districts of Lisbon and Porto.[11] In our empirical work below, we focus on the effect of competition on prices and treat a municipality, the area covered by a school license, as the relevant market area within which

[9]The age of a school's driving fleet was determined by mapping its vehicles' license plates to the year of initial registration using a key published by the Associação Automóvel de Portugal.

[10]Much of these data derive from Portugal's censuses of 2001 and 2011.

[11]The other excluded districts are Braga, Viana do Castelo, and Vila Real.

firms compete. While schools are required to conduct their business within the municipality covered by their license, students are free to frequent a school outside of their municipality of residence. Lisbon and Porto are large urban districts made up of densely populated municipalities in close proximity where the assumption of a single municipality comprising the relevant market area is likely unreasonable.

To collect price information in the remaining districts, we employed a team of 14 mystery shoppers who visited each school in person between November 2011 and March 2012. Each shopper used the same script to query school employees for price information, including both the upfront prices and the add-on surcharges. We include in our final sample all municipalities for whom our mystery shoppers were able to collect a complete fee schedule, consisting of the base course, the theory repetition, and the practice repetition, for all schools in the local market.[12] Our sample covers 150 municipalities with a total of 396 schools. These are displayed in panel (b) of Figure 1.

Table 1 contains descriptive statistics for our sample of markets. The average (median) municipality contains three (two) driving schools. They range in population from 3,445 to 160,825 residents, with an average population of 26,025 people, 13.8 percent of which are aged 15 to 24 years.

==========

Insert Table 1 about here

==========

In our descriptive analysis in Section 4, we incorporate information on these and a number of other market demographics. The lower portion of Table 1 summarizes these municipality-level demographics. The age and monthly wage income of the average resident is 40 years and €885.30 per month (for employed residents). The average illiteracy rate in our markets is 10.5 percent of the population aged 14 or higher, slightly above the national average of nine percent, and 35.2 percent of residents have completed at least compulsory secondary education, compared to 38 percent nationwide. A little more than half of all residents use cars on a daily basis, spending approximately 20 minutes in commuting time on average.

The price for the base course typically covers the exam center fees for both exams, all the requisite materials for the theory exam, and the required number of driving lessons from one of the school's instructors. If the student fails either exam, the student pays an additional

[12] We were forced to drop some municipalities where at least one school quoted unusually high prices for one of the two add-on services or was unable or unwilling to reveal the prices of the add-on services. We could not discern any systematic patterns in the frequency of such instances across districts or the territories assigned to each of our mystery shoppers. In total, we dropped 39 municipalities out of the 189 municipalities with at least one school in our subset of districts.

fee to cover the next exam center fee; the practice exam add-on price furthermore covers the required additional driving lessons. Since schools differ somewhat in what is included in the base course price, we standardize prices to be comparable across schools. We use the full price of the standard course of instruction as our price measure for all schools.

Panel (a) of Table 2 summarizes the distribution of prices across schools and markets. The median school charges €750 for its base driving course, with an interquartile range of €130. There is thus significant variation in prices. A large share of this variation is likely due to cost and demand differences across municipalities. Accordingly, the between-municipality standard deviation is 1.73 times the within-municipality standard deviation in upfront prices. The average practice exam add-on surcharge of €275.5 is more than double that of the theory exam add-on, €133.6, driven by the fact that the practice exam add-on includes additional lessons.

==========

Insert Table 2 about here

==========

3.2.2 Marginal Cost Estimation

Our model makes predictions for variation in markups across the upfront and add-on markets. To test these predictions, we require estimates of markups and thus marginal costs. An advantage to our setting is the relative simplicity of the service offered, which allows us to construct reliable estimates of the marginal cost a school incurs for serving an additional student in the main driving course, as well as in the repeat markets. In constructing our measures, we exploit information contained in a template that the industry association Associação Nacional dos Industriais do Ensino da Condução Automóvel (ANIECA) provides to new driving schools to estimate annual costs of operation, including both total and unit costs.

We use the cost template to compute marginal costs for each of the offered services: the upfront and the add-on services. The upfront service per student marginal cost consists of several individual components, including (i) the fees paid to the exam center for one theory and one practice test administration, (ii) the cost of instructional materials for the theory lessons, (iii) the instructor salary for 32 driving lessons and for the final on-road driving exam, and (iv) the vehicle operating costs of driving one of the school's vehicles during the practice lessons. According to ANIECA, vehicle operating costs consist of the cost of gasoline, depreciation expenses, maintenance, repair, and other expenses, and tolls and taxes. The theory repeat service entails the minimal marginal cost to the school of having to cover

the student's renewed theory exam fee, which is identical to the fee it pays to the exam center the first time the student takes the exam. Lastly, in offering the practice repeat course, the school incurs driving exam administration fees, as well as labor and vehicle maintenance costs equivalent to five additional driving lessons.

We use information on exam center fees, municipality-level salaries scaled to fall within the range of typical instructor pay, local gasoline prices, the estimated distance traveled in kilometers (km) during a 32-lesson and a five-lesson course of instruction, as well as the aggregate annual usage in km of each school's fleet of cars to derive marginal cost estimates. In the Appendix, we describe the procedure we use to calculate each of the cost components. Our detailed data allows us to recognize both spatial and firm-level variation in cost.

Here, we briefly summarize the magnitude of the different cost sources. Per student, the average school pays €58.45 in exam fees to cover the cost of one theory and one practice exam. Some amount of variation reflects differences in fees charged by private and publicly run exam centers. The cost of instructional materials is minimal and standardized, amounting to €10. We estimate that the school's labor cost of an instructor amounts to €237.15 and €43.11 for driving lessons during the upfront and practice repeat driving courses, respectively; we assume that no marginal labor costs arise from the in-class time spent by the instructor due to the common spare capacity in classroom space. The standard deviation in base course labor cost of €18.02 captures cost-of-living differences across municipalities reflected in municipality-level incomes. Vehicle operating costs represent the largest source of marginal cost with gasoline costs of €0.10 per km, estimated depreciation costs of €0.12 per km, maintenance, repair, and other expenses of €0.20 per km, and toll and road taxes of €0.03 per km.[13] When scaled by the 723 km that the average student covers during the driving course and exam, the marginal vehicle operating cost for the base course of instruction amounts to €327.31 on average. For the practice repeat course, the equivalent figure is €55.69.

Panel (b) of Table 2 summarizes the sum of the marginal cost components and the aggregate marginal cost of serving each student in the upfront and repeat driving courses. The average school incurs a marginal cost of €632.91 per student in the base course, with a standard deviation of €48.71. As a robustness check, we computed a simplified marginal cost measure that assumes constant annual usage and characteristics of the schools' fleet, which results in a similar average marginal cost of €628.35 for the base course. See Appendix Table A-1. We further verified the reliability of our cost estimates using feasibility studies that

[13] In total, we thus estimate a vehicle operating cost of €0.45 per km. This compares to estimates of vehicle operating and ownership costs provided by ACP (Automóvel Club de Portugal), Portugal's equivalent of AAA in the USA.

new school entrants prepare for the IMTT as part of the licensing process. These allowed us to compare our marginal cost estimates to the schools' own estimates for a country-wide sample of schools; the comparison suggests our estimates are reasonable. The add-on services generate estimated marginal costs of €17.43 and €139.81 for the theory and practice repeat courses, respectively.

3.2.3 Other School Characteristics

Here, we summarize other school characteristics, the size of the schools' student body, and market shares for the schools in our sample, and for the universe of driving schools in general. Table 2 summarizes the most pertinent information that we observe about schools.

Over the three-year period from 2008 to 2010, the average of our 396 sample schools enrolled 452 students, or approximately 150 students per year.[14] There is significant variation in enrollment figures across schools, however; the interquartile range of enrollment spans from 270 to 570 students. The number of a school's local competitors ranges from zero to 15 schools, with a typical municipality containing 3.0 schools. We use enrollment at the school and in the municipality to measure market shares. The average school's market share is 35 percent; the presence of monopoly markets skews the market share distribution with the median market share amountingt to 24 percent.

The median school's age is 13.2 years at the end of 2010, which reflects that approximately half of all sampled schools entered the market after deregulation in 1998. The schools have a driving fleet consisting of, at the median, three passenger vehicles, for which we observe characteristics such as an average displacement of 1.5 liters and an age of 5.3 years. Their median of five instructors have an average of 8.0 years of teaching experience at the school. Twenty-seven percent of schools have a functional website.

Furthermore, we record the straight-line distance of the school to the nearest district-wide IMTT office and to its most used exam center, to proxy for costs of interacting with IMTT and transporting students to the exam center. The average school is 28.4 km from the closest IMTT office. It also is 24.9 km from its most commonly used exam center; 67 percent of schools use at least one private exam center.

The schools in the pricing sample are representative of schools nationwide. Since we exclude the two urban districts of Lisbon and Porto, the sample schools are slightly smaller than the average school nationwide. Our markets are typically also more concentrated; as a result, the sample schools control a larger share of the market, relative to the average school's market share of 24 percent. The remaining school attributes, however, including the

[14] We employ data over a three-year period to smooth large fluctuations in the numbers of students for a small group of schools that faced such events as temporary closures due to repairs, and so on.

characteristics of the fleet and instructors are very similar.

3.2.4 Student Attributes

There are 54,178 unique learner driver candidates in our data. We only follow candidates for whom we have a complete history profile from enrollment at a school to completion of the practice exam, which eliminates approximately 27 percent of the full data set due to left or right truncation. The profile of an average student in our sample is representative of the entire Portuguese population of student drivers, however.

The demographic attributes that we observe for the students are limited. The average (median) student in our sample is 21.7 years (19.1 years) old. There is an even split of male and female students: 50.7 percent of our sample are females. We further observe the detailed location of the student's residence based on his seven-digit postal code, which designates approximately a city block. We calculate the straight-line distance from the postal code to the school that the student has chosen to attend. According to this measure, a majority of students are located less than three kilometers to their schools, a fact that alludes to the importance of distance to students' consideration and choice of schools. See Table 3. Clearly, in addition to competition from schools located in the same municipality, spatial differentiation is an important dimension that contributes to the variation in prices we see in the data.

==========

Insert Table 3 about here

==========

4 Pricing in the Upfront and Add-on Markets

Our theoretical model implies that the retake market exhibits monopoly- or collusive-like markups and that the upfront market exhibits competitive markups that decline in the number of market competitors. In what follows, we provide suggestive evidence that supports these model predictions.

While repeat courses are clearly additional services that add to the student's cost of driving instruction in the spirit of the add-on products studied by Ellison (2005) and others, their economic relevance would be minimal unless they were to generate significant profit for schools. Our model predicts that schools earn substantial markups from add-on services, and a casual comparison of prices and marginal costs in Table 2 suggests this to be the case. If few students actually require add-on services, however, then there is only limited scope

for substantive profits. We therefore begin by summarizing students' exam outcomes and school pass rates.

4.1 Incidence of Exam Repetitions

Figure 2 and Table 3 quickly dispel any notion that the add-on market is small. Among 54,178 students from 396 schools in our sample, 76.2 and 70.6 percent pass the theory and practice exams on the first attempt, respectively. More importantly, however, only 54.4 percent pass both exams on the first attempt. The average student takes the theory and practice exams approximately 1.35 and 1.41 times, respectively. This results in the driving school process being lengthy; the median student takes around 7.5 months from start to finish. Given that just under half of all students retake at least one exam, the retake market has the potential to serve as a significant source of profit for schools.

==========

Insert Figure 2 about here

==========

The data also provide evidence to support our assumption above that students are locked into the school by their initial choice of where to take the base course and that any exam repeat course can be treated as an unavoidable add-on to the base course: students rarely switch schools. Only 1.8 and 0.1 percent of all students transfer schools during the theory and practice exam phases, respectively, while only 0.7 percent transfer in between the theory and practice exam phases. The majority of these, or 65 percent, transfer to schools outside their original school's municipality, suggesting that exogenous reasons such as moving explain a significant share of transfers. Schools engage in several practices that make it prohibitively costly for students to transfer: first, the fee for the upfront course of instruction is non-refundable, making it unlikely that students switch before reaching their first practice exam. Transferring schools further requires restarting the base course from the beginning at a new school. Schools price their repeat courses below their own and their competitors' base courses, so that students who consider transferring schools comparison shop their current schools' add-on prices to competitor schools' relatively higher base prices.

4.2 Markups

To assess the per-student profitability of offering the three services, we consider two common measures of profit margins: the traditional markup or price-cost margin, and the percentage markup (or Lerner index), which equals the ratio of markup to price. Table 4 reports a

summary of markups across schools; Figure 3 illustrates the distribution of markups. The markups for the add-on services are higher on average both in levels and in percentage terms than those of the upfront service.

==========

Insert Table 3 about here

==========

The average Lerner index for the upfront service is 13.6 percent. Were firms fully homogeneous and competing in quantities, a market with approximately seven firms would produce a Lerner index of this magnitude. In contrast, the average market in our data has only three firms. This suggests that, in order to justify the Lerner indices we observe, the upfront market must be more competitive than we would otherwise predict by simply using the number of schools in the market as a proxy of competitiveness. In support of our thesis, the percentage markups furthermore reveal extraordinary market power in the add-on markets, with averages of 85.8 percent and 47.7 percent in the theory and practice add-on markets, respectively.

Moreover, the variation in markups differs significantly across the primary and secondary markets. If schools act according to the stylized model above, we might expect there to be limited variation in add-on markups to reflect that schools charge the highest sustainable price in the market. The standard deviation of upfront markups across markets is €110.25, or 13.1 percent for the Lerner index, while the standard deviation of theory and practice add-on markups across markets is €29.40 (7.8 percent for the Lerner index) and €47.27 (10.6 percent), respectively.

Sixty-two schools, or 15.7 percent of the sample, have negative markups in the upfront market, pricing below estimated cost. Most of these schools garner positive effective markups across the three services, however; they are able to cover their total variable costs with the revenues they earn from the add-on markets despite the losses they incur in the upfront market. We compute effective markups as the total fees paid by each student less the total variable cost incurred in serving the student, averaging across students at each school. For every euro the average school earns in profit from offering the base course, it earns an additional €0.87 in effective markup on the full cost of serving its average student. At the student level, the average student in the sample yields an additional €0.66 for every €1 he generates in profit for the upfront service. Thus schools with a larger student population tend to have lower effective markups on average than smaller schools. According to this measure, a school earns €214.77, or 20.3% of fees paid by the student, on the average student.[15]

[15]We estimate that nine, or 2.2 percent of, schools have negative effective markups of on average −€48.45. We are unable to determine whether this is because of error in our marginal cost estimate or because these

Implicitly, our model posits that the upfront and add-on markets are distinct: the upfront market prices and markups depend on a number of consumer- and market-related factors, while the add-on market surcharges and markups are independent of these. Evidence points to little relationship between the upfront and add-on prices: only a weak association exists between the upfront prices and the theory exam add-on prices (0.046) and the practice exam add-on prices (0.037). On the other hand, there is a strong and positive association between each of the add-on prices (0.436). Similarly, for markups, the correlations between base markup and the theory and practice markups are 0.022 and 0.073, respectively, but 0.411 between the theory and practice markups themselves. These correlations provide indirect evidence that factors affecting the setting of schools' upfront prices differ from those affecting their add-on prices.

4.3 Sources of Profit

Perhaps most compelling to the add-on market profit motive is that a significant proportion of schools' variable profits derive from add-on fees. The average (median) school derives 11.6 percent (12.0 percent) and 18.7 percent (16.7 percent) of total variable profits from the theory and practice exam add-on services, respectively, or 30.3 percent (28.8 percent) of profit overall. At the same time, schools earn a smaller share of total revenues from add-on services: the average (median) school derives 4.4 percent (4.3 percent) and 9.5 percent (8.9 percent) of total revenues from the two add-on services for a total of 13.9 percent (13.4 percent) of revenue. This suggests that schools face relatively lower variable costs for add-ons compared to the cost of the corresponding base course, and, thus, more revenue is passed on as variable profits to schools in the add-on market.

Figure 4 reveals significant variation in the percentage of total variable profits derived from the add-on services. Eighty schools (20.2 percent of the sample) earn more than 20 percent of their variable profits from the theory exam add-on service, with 28 of these earning more than 30 percent from theory exam retakes. The interquartile range for the share of profits stemming from the theory exam retake market is 8.1 to 17.6 percent. Typically, the share of profit from practice exam retakes further exceeds these magnitudes with an interquartile range of 10.5 to 23.8 percent. 148 schools (or 37.4 percent of the sample) earn more than 20 percent of their variable profits from the practice exam add-on service, with 42 and 23 of these schools earning more than 30 and 40 percent from practice exam retakes, respectively. The variation in the two add-on markets translates into significant variation in the share of profit stemming from add-ons in aggregate, with an interquartile range of 21.4

schools are truly making negative profit.

to 37.8 percent.

==========

Insert Figure 4 about here

==========

4.4 Determinants of Markups

The previous section establishes that schools have a profit motive in the add-on market with significantly higher markups being earned than on the base course. We now test whether the observed markups for the upfront and retake services are consistent with the basic predictions of the bargains-then-ripoffs model from Section 2. The model predicts that variation in markups on the upfront base service, but not the add-on surcharges, depend on the number of competitors in the market, on the proportion of myopes, and on the type-specific probability of failing the exam. We consider a number of reduced form models of markups to test these predictions, while controlling for heterogeneity in demand conditions across schools that might introduce variation in prices and markups unrelated to the competitive environment.

We regress school markups on the number of competitors within the municipality as well as a number of other school- and municipality-related controls. While we benefit from an accurate measure of the number of market competitors and the overall probability of failing on both exams, we do not have separate proxies for the proportion of myopes and the probability of failing by consumer type. As an extension to the current work, we plan to estimate an empirical version of the demand model above, using the consumers' responses to variation in upfront and base prices to pin down the share of myopic consumers consistent with the observed school choices. A direct benefit to estimating such a model is that it allows us to derive estimates of variation in consumer types across markets and demographic groups.

In our regression analysis, we employ both OLS and two-stage least squares techniques to control for the possibility that the number of schools in the municipality are endogenous to schools' markups. As discussed in Section 3.1, the driving school industry faced municipality-based entry restraints prior to being deregulated in 1998. These restraints were closely tied to the municipality's population, stipulating that each licensed school serves a minimum of 25,000 residents in the chosen municipality or is the sole provider in a municipality with less than 25,000 residents. The regulated number of schools in each municipality prior to liberalization thus serves as a good instrument for the number of firms: by being tied to population alone, it is likely uncorrelated with unobserved profit shifters that are reflected in both markups and entry in the unregulated regime, but is related to the number of schools

in the municipality under free entry, post deregulation.[16]

The results of the regressions of the upfront profit margins on the number of market competitors with school- and municipality-level controls are shown in Table 5. While the first-stage F-statistics and partial R-squared statistics for the instrumental variables regressions imply a strong instrument, the endogeneity concerns do not seem to drive the results, which themselves appear to be robust across all specifications (including the specification that incorporates the share of the population under the age of 25 as an instrument for the number of market competitors). Similarly, the introduction of a number of school- and municipality-related controls do not change the coefficient on the number of market competitors significantly. Across specifications, the number of schools in the municipality has a strong and significantly negative impact on the upfront markups of schools in that municipality. For every school added to the market, the upfront markup decreases by approximately €23.38, corresponding to 25.8 percent of the average upfront markup. This result aligns with the theoretical prediction that upfront markups should fall in the number of competitors because of the highly competitive nature of the upfront market.

==========

Insert Table 5 about here

==========

The results for the equivalent regressions of the add-on profit margins on the number of market competitors with school- and municipality-level controls are shown in Table 6. As with the upfront markup regressions, the results are robust to instrumenting for the number of firms and introducing school- and municipality-level controls. In each of these specifications, the number of schools in a school's own municipality is statistically insignificant in affecting both add-on markups and has an economically much smaller effect on both markups for each additionally competitor. These results are thus suggestive of there being no competitive influence on either retake markup, with schools benefitting from highly localized market power.

==========

Insert Table 6 about here

==========

[16] As a robustness check (not shown) we further incorporated the share of the population above the age of 15 but below the age of 25 as an instrument for the number of firms: the vast majority of students in our sample (84.2%) are below age 25 and the size of that age category in the municipality thus represents a good proxy for the size of the municipality's market. (Note that while students need to at least be of age 18 to enroll in driving school, we were not able to obtain data for finer age bins from Statistics Portugal.) While the market size shifts total variable profit in the market, and thus the profitability of entering, it should not affect the markup charged per student if marginal costs are near constant and economies of scale are small.

Note that the endogenous variable is a discrete count of the number of schools in the market, but our linear regression models treat this regressor as continuous, and furthermore assume that the effect of an incremental entrant on markups remains constant regardless of how many schools end up in the market. We investigate the sensitivity of the results to these assumptions by estimating a specification that models the number of schools as an ordered probit model, allowing the unobservable market-level error that drives entry decisions to be correlated with an unobserved market-level error that affects markups. This correlation represents the possible correlations between unobserved market structure shifters and unobserved markup determinants that we worry might bias the estimated effect of the number of schools on prices. We estimate the resulting nonlinear two-equation model using FIML and derive the likelihood in the Appendix.

One market where we might expect higher upfront markups, but similar add-on markups compared to other markets, is a monopoly market. Fifty-three schools, or approximately 13.4 percent of sampled schools, serve as the only school in their municipality; these schools should wield monopoly power in both upfront and add-on markets. In fact, the average (median) school in a monopoly market earns €153.49 (€160.03) in upfront markups, compared to €107.35 (€105.08) earned by the average (median) school in an oligopoly market (p-value of test of the equality of means of 0.0054). Schools in both monopoly and oligopoly markets earn about the same markups in the add-on market, however: the average (median) school in a monopoly market garners €119.65 (€124.00) and €135.17 (€133.25) in theory and practice add-on markups, respectively, while the average (median) school in a oligopoly market receives €115.61 (€119.00) and €135.71 (€138.95) in theory and practice add-on markups, respectively, and we cannot reject the hypothesis of equal means, with p-values of 0.2918 and 0.9400, respectively. The similarity of add-on markets across monopoly and oligopoly markets, combined with the demonstrated decline in markups as the number of competitors rises, in particular from monopoly to oligopoly markets, thus bear out the model predictions of loss-leader pricing.

5 Conclusion

In this paper we have modeled a market in which firms take advantage of locked-in myopic consumers by charging high surcharges in unavoidable add-on markets. When consumers have limited foresight about their future demand for a product and face nearly infinite switching costs, firms have the dual strategic incentive to charge as high a price as the add-on market can support and to charge a correspondingly low price in the upfront market to entice consumers to their firm in the first place via seemingly low "loss leader" prices.

Our model, hence, predicts upfront (add-on) markups that (do not) vary with the market structure a firm faces. The model also implies that all markups – both upfront and add-on – depend on the proportion of consumer types in the market and the type-specific explicit probability of requiring the unavoidable add-on.

We present evidence of these phenomena in the context of Portuguese driving schools. With data on prices and constructed marginal costs we demonstrate that schools not only face a strong profit motive for setting high surcharges in the add-on market, but also earn high, monopoly- or collusive-like markups in the add-on markets while earning relatively low markups in the upfront market. The evidence corroborates our model predictions that the latter vary with the level of competition whereas add-on markups do not.

Our research benefits from a highly detailed data snapshot of the driving school industry, which provides us with an opportunity to answer additional questions of interest, most notably who the typical myopic consumer is and which consumers participate most in the add-on market, implicitly subsidizing others. As a next step, therefore, we plan to estimate an empirical model of school choice. Since we believe there to be two natural segments in the market comprised of myopes and sophisticates, we specify a modified latent class model of demand with the intent of empirically pinning down the proportion of the two consumer types consistent with the observed school choices under schools' chosen base and add-on prices. Evidence that speaks to the question of who subsidizes whom in market such as ours is of significant normative policy interest in considering regulations concerning add-on pricing, which in the case of policies under consideration by the IMTT range from requiring schools to inform consumers about typical propensities of failing exams to directly or indirectly regulating prices in the add-on market.

References

Beggs, A. W. and Klemperer, P. (1992). Multi-period competition with switching costs, Econometrica **60**(3): 651–66.

Bertini, M. and Wathieu, L. (2008). Attention arousal through price partitioning, Marketing Science **27**(2): 236–246.

Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E. and Zinman, J. (2010). What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment, Quarterly Journal of Economics **125**(1): 263–306.

Brown, J., Hossain, T. and Morgan, J. (2010). Shrouded attributes and information suppression: Evidence from the field, **125**(2): 859–876.

Chetty, R., Looney, A. and Kroft, K. (2009). Salience and taxation: Theory and evidence, American Economic Review **99**(4): 1145–77.

Chevalier, J. A., Kashyap, A. K. and Rossi, P. E. (2003). Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data, American Economic Review **93**(1): 15–37.

DellaVigna, S. and Malmendier, U. (2004). Contract Design and Self-Control: Theory and Evidence, The Quarterly Journal of Economics **119**(2): 353–402.

DellaVigna, S. and Malmendier, U. (2006). Paying not to go to the gym, American Economic Review **96**(3): 694–719.

Diamond, P. A. (1971). A model of price adjustment, Journal of Economic Theory **3**(2): 156–168.

Ellison, G. (2005). A model of add-on pricing, The Quarterly Journal of Economics **120**(2): 585–637.

Ellison, G. and Ellison, S. F. (2009). Search, obfuscation, and price elasticities on the internet, Econometrica **77**(2): 427–452.

Farrell, J. and Klemperer, P. (2007). Coordination and lock-in: Competition with switching costs and network effects, in M. Armstrong and R. H. Porter (eds), Handbook of Industrial Organization Vol. III, North-Holland, Amsterdam.

Gabaix, X. and Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets, The Quarterly Journal of Economics **121**(2): 505–540.

Grubb, M. D. (2009). Selling to overconfident consumers, American Economic Review **99**(5): 1770–1807.

Heidhues, P. and Köszegi, B. (2010). Exploiting naïveté about self-control in the credit market, American Economic Review **100**(5): 2279–2303.

Heidhues, Paul, B. K. and Murooka, T. (2012). The market for deceptive products. Working Paper, University of California Berkeley.

Holton, R. H. (1957). Price discrimination at retail: The supermarket case, The Journal of Industrial Economics **6**(1): 13–32.

Hossain, T. and Morgan, J. (2006). ...plus shipping and handling: Revenue (non) equivalence in field experiments on ebay, The B.E. Journal of Economic Analysis & Policy **6**(2): 3.

Jin, G. Z. and Leslie, P. (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards, Quarterly Journal of Economics **118**(2): 409–451.

Jovanovic, B. (1982). Truthful disclosure of information, Bell Journal of Economics **13**(1): 36–44.

Klemperer, P. (1987a). Markets with consumer switching costs, The Quarterly Journal of Economics **102**(2): 375–94.

Klemperer, P. (1987b). The competitiveness of markets with switching costs, RAND Journal of Economics **18**(1): 138–150.

Kosfeld, M. and Schüwer, U. (2011). Add-on pricing, naive consumers, and the hidden welfare costs of education. IZA Discussion Paper No. 6061.

Köszegi, B. and Rabin, M. (2006). A model of reference-dependent preferences, Quarterly Journal of Economics **121**(4): 1133–1165.

Lal, R. and Matutes, C. (1994). Retail pricing and advertising strategies, The Journal of Business **67**(3): 345–370.

Miao, C.-H. (2010). Consumer myopia, standardization and aftermarket monopolization, European Economic Review **54**(7): 931–946.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications, Bell Journal of Economics **12**(2): 380–391.

Morwitz, V. G., Greenleaf, E. A. and Johnson, E. J. (1998). Divide and prosper: Consumers' reactions to partitioned prices, Journal of Marketing Research **35**(4): 453–463.

Office of Fair Trading (2010). Market study report: Advertising of prices. Available at `http://www.oft.gov.uk/OFTwork/markets-work/advertising-prices/`, accessed 5/30/2012.

Oster, S. M. and Scott Morton, F. M. (2005). Behavioral biases meet the market: The case of magazine subscription prices, The B.E. Journal of Economic Analysis & Policy **5**(1): 1.

Spiegler, R. (2006). The market for quacks, Review of Economic Studies **73**(4): 1113–1131.

Spiegler, R. (2011). Bounded Rationality and Industrial Organization, Oxford University Press.

Stango, V. and Zinman, J. (2009). What Do Consumers Really Pay on Their Checking and Credit Card Accounts? Explicit, Implicit, and Avoidable Costs, American Economic Review **99**(2): 424–429.

Stole, L. (2007). Price discrimination and imperfect competition, in M. Armstrong and R. H. Porter (eds), Handbook of Industrial Organization Vol. III, North-Holland, Amsterdam.

Verboven, F. (1999). Product line rivalry and market segmentation–with an application to automobile optional engine pricing, Journal of Industrial Economics **47**(4): 399–425.

Walters, R. (1988). Retail promotions and retail store performance: A test of some key hypotheses, Journal of Retailing **64**(2): 153–180.

Wertenbroch, K. (1998). Consumption self-control by rationing purchase quantities of virtue and vice, Marketing Science **17**(4): 317–337.

Figures and Tables

Figure 1: Driving schools by municipality, mainland Portugal

(a) All municipalities

(b) Available sample of municipalities

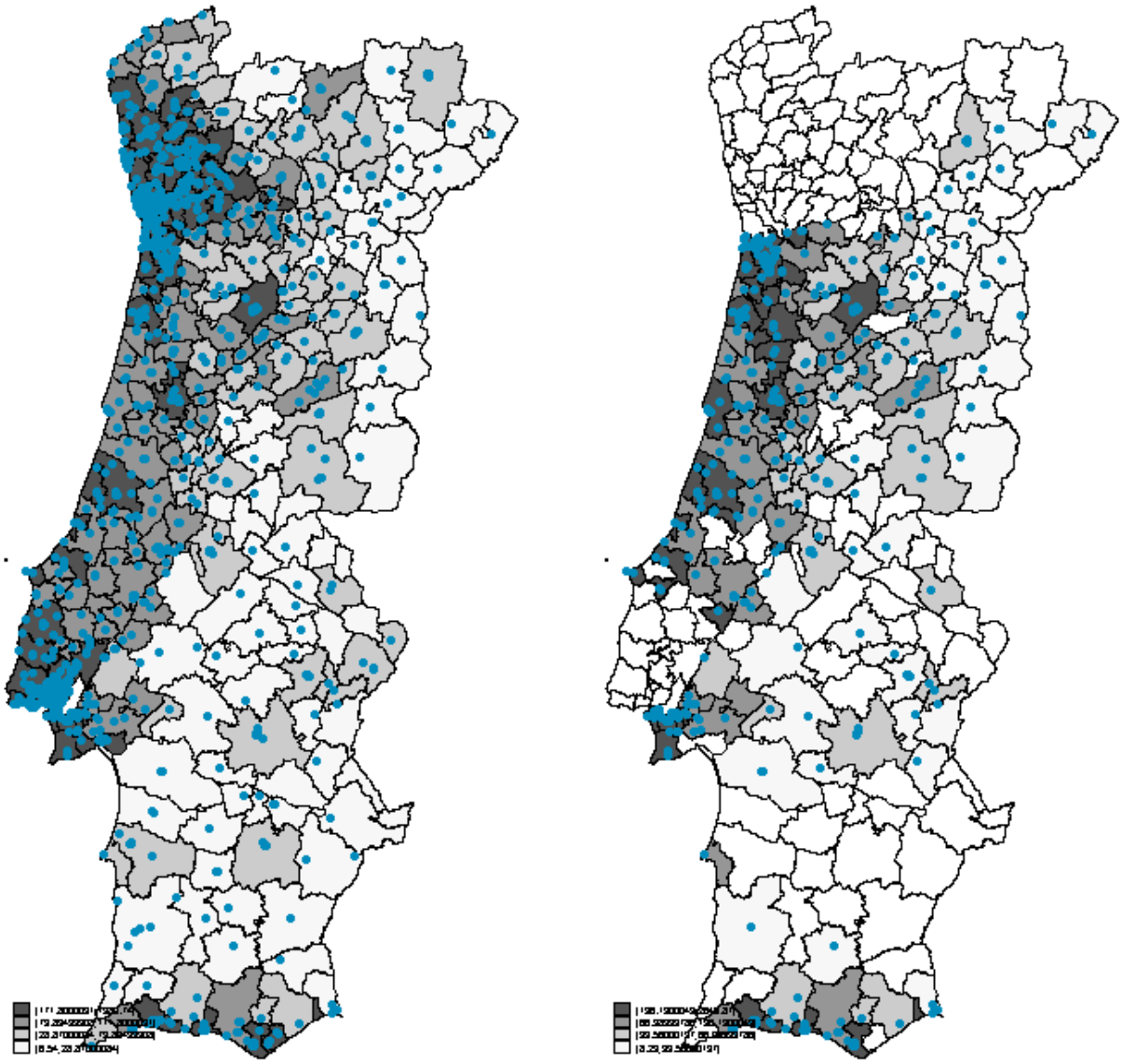


Figure 2: Number of practice and theory exams taken, by student

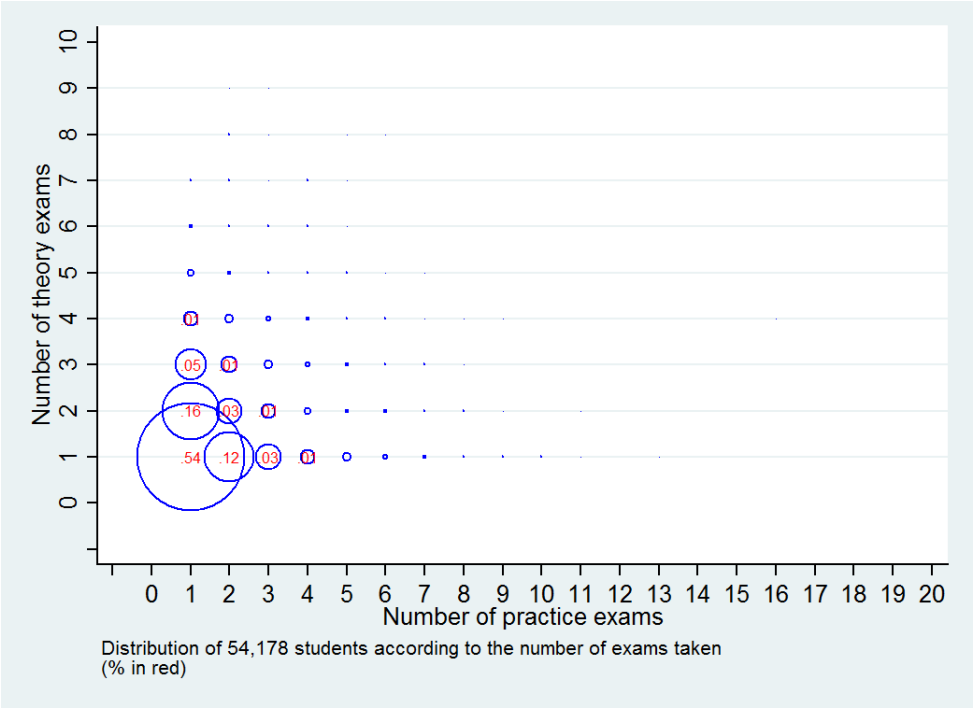
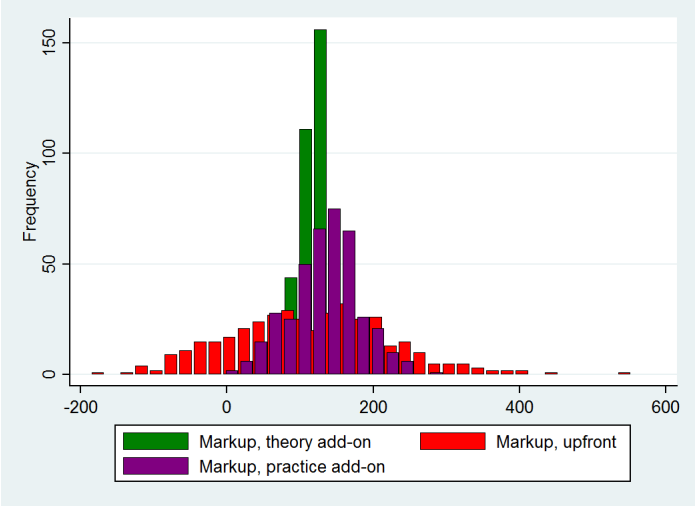


Figure 3: Distribution of markups in upfront and add-on markets, by school

(a) Markup ($)



(b) Markup/Price (%)

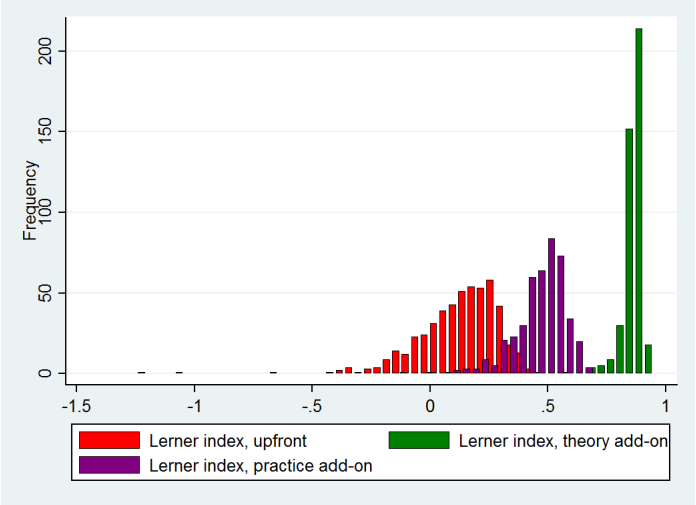
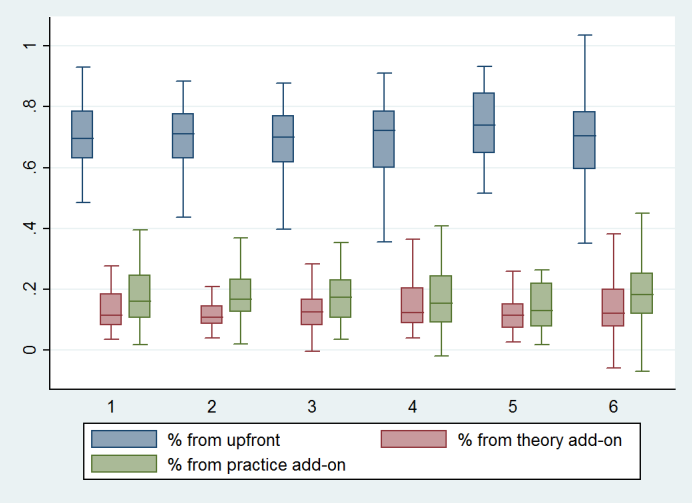


Figure 4: Variable profit from base and add-on products, by number of competitors



Note: we combine the 22 markets with six or more competitors into the category labeled 6.

Table 1: Summary Statistics, Municipality Characteristics, Sample Markets

	Mean	Std Dev	Q25	Med	Q75
Number of schools	2.99	2.68	1	2	4
Population (ppl)	26,024.59	28,406.34	8,585	15,506	31,482
Population aged 15-24 (%)	13.80	1.10	13.22	13.72	14.64
Population density (ppl/km)	559.27	1061.21	44.14	114.33	470.42
Mean resident age (yr)	40.31	2.79	38.16	40.20	41.85
Mean per-capita income (€)	885.30	112.96	822.20	882.70	940.10
Unemployment rate (%)	6.50	1.94	5.10	6.40	7.90
Secondary education completed (%)	35.20	9.82	28.41	34.41	40.53
Illiteracy rate (%)	10.49	4.48	6.95	9.72	12.27
Daily vehicle usage (%)	52.75	7.95	47.03	54.46	58.43
Mean commute time (min)	19.56	7.34	15.19	16.22	20.69
Vehicles sold (no/1000 ppl)	15.13	6.38	10.58	14.56	19.21
Gasoline price (€/L)	1.42	0.04	1.39	1.41	1.44

Note: We calculate the population density for each parish in a given municipality and report a weighted average population density. Statistics other than number of schools and population are population weighted. Mean income refers to average monthly wage income of full-time employees.

Table 2: Summary Statistics, School Characteristics

	Sample Schools ($N = 396$)					All Schools ($N = 1,141$)				
	Mean	Std Dev	Q25	Med	Q75	Mean	Std Dev	Q25	Med	Q75
(a) Prices										
Upfront	746.43	106.93	670	750	800					
Theory add-on	133.58	29.72	120	140	150					
Practice add-on	275.45	47.02	250	280	300					
(b) Marginal Cost										
Upfront	632.91	48.71	594.30	632.53	668.59					
Theory add-on	17.43	2.32	15	17	21					
Practice add-on	139.81	12.46	130.93	139.84	149.25					
(c) Other Characteristics										
Number of students (3yr)	452.31	283.94	270	379	570	492.43	345.25	280	419	608
Market share (%)	34.89	30.40	12.27	23.56	46.22	23.80	26.85	6.13	12.86	30.22
Age (yrs in 2010)	21.77	16.46	8.03	13.23	30.74	20.27	16.09	8.29	10.41	29.46
Website? (1=Y,0=N)	0.27	0.44	0	0	1	0.28	0.45	0	0	1
Number of instructors	5.40	3.46	3	5	7	4.90	3.29	3	4	6
Instructor experience (yr)	8.00	5.85	3.79	6.17	10.23	7.72	5.63	3.88	6.12	9.89
Number of vehicles	3.48	1.91	2	3	4	3.56	2.38	2	3	4
Mean vehicle age (yr)	5.32	3.51	2.48	4.34	7.65	5.88	3.59	3.00	5.13	8.58
Median vehicle displacement (L)	1.47	0.15	1.38	1.46	1.49	1.49	0.17	1.37	1.46	1.53
Distance to exam center (km)	24.90	19.64	8.55	21.45	35.75	20.21	18.65	6.34	15.33	28.38
Private center option? (1=Y,0=N)	0.67	0.47	0	1	1	0.72	0.45	0	1	1
Distance to IMTT office (km)	28.42	17.86	15.03	27.36	43.21	23.35	17.83	8.55	20.24	35.19

Table 3: Summary Statistics, Student Attributes ($N = 54,178$)

			Distribution				
	Mean	Std Dev	Min	Q25	Med	Q75	Max
Age at theory exam (yr)	21.73	6.40	17.67	18.38	19.12	21.80	76.45
Gender (1=F,0=M)	0.51	0.50	0	0	1	1	1
Distance to school (km)	4.90	6.06	0	0.96	2.82	6.41	39.98
Theory exam taken (no)	1.35	0.76	1	1	1	1	16
Practice exam taken (no)	1.41	0.75	1	1	1	2	9
Pass rate, first theory exam (%)	0.77	0.42	0	1	1	1	1
Pass rate, first practice exam (%)	0.71	0.46	0	0	1	1	1
Time to completion (days)	263.35	152.66	16	154	224	332	1,081

Table 4: Markups for upfront and add-on services

Markups (€)							
			Distribution				
Service	Mean	Std Dev	Min	Q25	Med	Q75	Max
Upfront	113.52	110.25	−186.54	33.99	111.05	187.37	541.46
Theory add-on	116.15	29.40	−2.00	103.00	119.00	133.00	229.00
Practice add-on	135.64	47.27	−3.09	109.21	137.72	165.16	290.26
Effective, all services	214.77	124.28	−149.66	127.03	216.13	304.54	602.08
Percentage Markup (%)							
Upfront	13.63	13.12	−33.31	5.34	15.28	23.24	46.08
Theory add-on	85.83	7.79	−13.33	85.00	86.92	88.67	92.68
Practice add-on	47.67	10.61	−2.47	42.64	49.78	54.52	68.26
Effective, all services	20.31	11.48	−25.02	12.63	21.85	28.75	47.56

The row labeled "effective, all services" refers to the effective markup the firm earns across the upfront and add-on markets, as defined in the text.

Table 5: Regression Models of Upfront Price Markups

	Dependent Variable: Upfront Price Markups (€) ($N = 420$)			
	OLS		IV	
Variable	(i)	(ii)	(i)	(ii)
Number of competitors	−18.4683***		−23.0808***	
	(5.2550)		(6.1360)	
% fail on both exams	−184.5350**	−183.6603**	−189.3607**	−171.6006*
	(93.4410)	(93.5556)	(94.3611)	(93.4605)
$I_{\{N=2\}}$		−28.5762		−40.3795*
		(19.9820)		(20.6757)
$I_{\{N=3\}}$		−39.7401*		−43.8094**
		(23.7688)		(21.3096)
$I_{\{N=4\}}$		−68.7008**		−99.5310***
		(29.7309)		(28.7616)
$I_{\{N=5\}}$		−84.1256***		−101.5144***
		(31.6220)		(26.3003)
$I_{\{N\geq 6\}}$		−89.2456***		−137.2424***
		(28.2628)		(26.2560)
Adjusted R^2	0.2290	0.2270	——	
1st Stage Partial R^2	——	——	0.7907	
1st Stage F-Statistic	——	——	532.1709	
Log-Likelihood (FIML)	——	——	——	2,520.74
Instrument(s)	——	——	Pre-deregulation nº schools	Pre-deregulation nº schools

We include as controls each school's median age of its driving fleet and indicator variables for whether each school's competitors are between 5km and 10km away or are farther than 10km away, each school's municipality-level per capita purchasing power index and dynamic response purchasing power factor, and each school's parish-level mean rental rates. Across the first four specifications, the average F-statistic for the joint significance of the controls equals 26.5700 (p-value 0.0001). In the final specification we drop the monopoly market indicator variable for the sake of identification.

Table 6: Regression Models of Add-on Price Markups

	Dependent Variable: Add-on Prices Markups (€) ($N = 420$)			
	OLS		IV	
Variable	(i)	(ii)	(i)	(ii)
Number of competitors	4.1854		3.3407	
	(3.1951)		(3.7721)	
% fail on both exams	−15.7688	−39.8110	−12.0822	−56.0758
	(73.0732)	(72.7070)	(73.0986)	(71.4941)
$I_{\{N=2\}}$		13.1085		12.2008
		(12.0279)		(12.8591)
$I_{\{N=3\}}$		0.3514		−3.5648
		(13.9764)		(14.3474)
$I_{\{N=4\}}$		−20.5619		−25.6160
		(16.6379)		(16.7341)
$I_{\{N=5\}}$		−10.7747		−10.3948
		(17.6024)		(16.6559)
$I_{\{N\geq 6\}}$		40.4319**		42.1233**
		(16.2740)		(18.4286)
Adjusted R^2	0.0580	0.1314	——	
1st Stage Partial R^2	——	——	0.7983	
1st Stage F-Statistic	——	——	587.1780	
Log-Likelihood (FIML)	——	——	——	2,386.53
Instrument(s)	——	——	Pre-deregulation nº schools	Pre-deregulation nº schools

For a list of included controls, see footnote to Table 5. Additionally, we include each school's municipality-level proportion of residents using public transportation. Across the first four specifications, the average F-statistic for the joint significance of the controls equals 10.0240 (p-value 0.0230). In the final specification we drop the monopoly market indicator variable for the sake of identification.

Appendix

A. Equilibrium in Illustrative Model in Section 2

Solution in the Absence of Shrouding

In this Appendix, we establish the equilibrium of the add-on pricing model in section 2. Here, we assume that the full menu of prices is observable to consumers; myopic students simply choose not to take add-on prices into account when making their school choice. In the following section we establish that in contrast to markets where the add-on is avoidable, there is no profit gain to the firm from shrouding add-on prices. Recall the properties of the equilibrium summarized in Equation 3:

Suppose there are n schools that offer an upfront service for price p_u at cost c_u and an add-on service for price p_a at cost c_a, and that there are a continuum of students. Let the fraction of myopes in the student population be $\alpha \in (0,1)$ and the fraction of students who fail the exam be $\overline{\lambda} \in (0,1]$. If students fail the exam, they must buy the add-on service in the second period. There is a unique symmetric Nash equilibrium in which sophisticates engage in effort to lower their probability of failing to $\underline{\lambda}$ and schools charge an upfront price $p_u^* = c_u + \frac{\sigma n}{n-1} - \left((1-\alpha)\overline{\lambda} + \alpha\underline{\lambda}\right)(\overline{p}_a - c_a)$ in period 1 and an add-on price $p_a^* = \overline{p}_a > c_a$ in period 2.

Proof. Consider first the school's choice of add-on price, p_a. In the second period, school j sells the add-on service to α sophisticated and $(1-\alpha)$ myopic students and earns profit of

$$\begin{aligned}\pi_j = {} & \alpha\left(p_{u,j} - c_u + \underline{\lambda}\left(p_{a,j} - c_a\right)\right) D^s\left(p_{u,-j} - p_{u,j} + \underline{\lambda}(p_{a,-j} - p_{a,j})\right) \\ & + (1-\alpha)\left(p_{u,j} - c_u + \overline{\lambda}\left(p_{a,j} - c_a\right)\right) D^m\left(p_{u,-j} - p_{u,j}\right),\end{aligned} \tag{A1}$$

where we express the school's demand, $\{D^m, D^s\}$, in equation 2 only as a function of the price arguments.

Given that the add-on price, p_a, does not shift the demand of the myopic students, it is optimal for the firm to set it at the highest possible level, $\overline{p}_a$. A lower add-on price would not be profit-maximizing: for any combination of competitor prices, $\{p_{u,-j}, p_{a,-j}\}$, the firm could raise its add-on price by Δ and lower its upfront price by $\underline{\lambda}\Delta$. This would leave the demand and per-student revenue earned on sophisticated students unchanged. It would, however, increase both the demand from myopic students through the decline in the upfront price and - provided $\overline{\lambda} > \underline{\lambda}$ - the revenue per myopic student.[17]

[17] If $\overline{\lambda} < \underline{\lambda}$, a similar argument results from raising the add-on price by Δ and lowering the upfront price

The firm's upfront price p_u then maximizes:

$$\pi_j = \left(p_{u,j} - c_u + \left(\alpha\underline{\lambda} + (1-\alpha)\overline{\lambda}\right)(\overline{p}_a - c_a)\right) D^m\left(p_{u,-j} - p_{u,j}\right). \tag{A2}$$

Since schools are symmetric and charge the same add-on price in equilibrium, relative differences in add-on prices do not affect the sophisticated students' school choice. As a result, at the optimal add-on prices, their demand equals the myopic students' demand D^m. Solving the first-order condition to the firm's upfront pricing problem results in equilibrium prices of:

$$p_u^* = c_u + \frac{\sigma n}{n-1} - \left((1-\alpha)\overline{\lambda} + \alpha\underline{\lambda}\right)(\overline{p}_a - c_a). \tag{A3}$$

This exposition assumes that sophisticated students find it in their best interest to engage in costly effort to reduce their probability of failing to $\underline{\lambda}$. They will do so provided the cost savings from doing so, $\left(\overline{\lambda} - \underline{\lambda}\right)\overline{p}_a$, exceed the cost of effort e. Otherwise, the optimal upfront price simplifies to:

$$p_u^* = c_u + \frac{\sigma n}{n-1} - \overline{\lambda}\left(\overline{p}_a - c_a\right). \tag{A4}$$

Equation A4 also illustrates the profit-neutrality inherent in the add-on pricing model with symmetric types. In equilibrium, with equal probabilities of failing across types, the firm earns expected revenue per student of

$$p_u^* + \overline{\lambda}\overline{p}_a = c_u + \overline{\lambda}c_a + \frac{\sigma n}{n-1}. \tag{A5}$$

The same level of revenue would result in a pricing game where firms served only sophisticated students who accounted for both the upfront and the add-on services in their school choice. Then, the firm's profit function would be:

$$\pi_j = \left(p_{u,j} + \overline{\lambda}p_{a,j} - \left(c_u + \overline{\lambda}c_a\right)\right) D^s\left(p_{u,-j} - p_{u,j} + \overline{\lambda}(p_{a,-j} - p_{a,j})\right) \tag{A6}$$

In the absence of myopic types, there is no unique solution to the firm's pricing problem to pin down p_u^* and p_a^* if firms commit to prices in the first period. Only the expected per-student revenue is uniquely identified; it equals to the expected revenue in Equation A5. Add-on pricing in the presence of myopic students thus does not change the total amount a school earns from a student in expectation. It does, however, pin down how that revenue is distributed over the two services, placing a higher monetary burden on exam repeaters.

by $\overline{\lambda}\Delta$.

Introduction of Shrouding

In this section, we consider possible profit gains to the firms from shrouding - that is not making public - their add-on prices. This choice does not affect myopes' demand; however, sophisticated consumers now make their school choice based on an expected price of the add-on product, Ep_a. Rational consumers realize that they are locked into their school by the time the add-on purchase decision is determined; as a result, they expect the add-on price to be $\overline{p}_a$ when encountering a school with shrouded prices.

Consider now the best response of a school whose $(n-1)$ competitors shroud their add-on price. If the school shrouds as well, sophisticated consumers face only schools with shrouded prices. They expect all add-on prices to equal $\overline{p}_a$ and do not rely on the add-on price as a determinant of school choice. As a result, school j's profit is given by:

$$\begin{aligned}\pi_j &= \alpha\left(p_{u,j} - c_u + \underline{\lambda}\left(p_{a,j} - c_a\right)\right) D^s\left(p_{u,-j} - p_{u,j}\right) \qquad \text{(A7)}\\ &\quad + (1-\alpha)\left(p_{u,j} - c_u + \overline{\lambda}\left(p_{a,j} - c_a\right)\right) D^m\left(p_{u,-j} - p_{u,j}\right)\\ &= \left(p_{u,j} - c_u + \left(\alpha\underline{\lambda} + (1-\alpha)\overline{\lambda}\right)\left(p_{a,j} - c_a\right)\right) D^m\left(p_{u,-j} - p_{u,j}\right) \qquad \text{(A8)}\end{aligned}$$

Given that demand does not depend on add-on prices, the firm indeed optimally sets the add-on price as high as possible, or at $\overline{p}_a$. The firm's choice of upfront price now again maximizes the profit given by equation A2, and the same set of optimal prices $\{p_u^*, p_a^*\}$ arises in a fully shrouded equilibrium as in the unshrouded equilibrium above.

Does a school have an incentive to unshroud its add-on price when its competitors shroud their add-on prices? In that case, the firm chooses its price pair to maximize:

$$\begin{aligned}\pi_j &= \alpha\left(p_{u,j} - c_u + \underline{\lambda}\left(p_{a,j} - c_a\right)\right) D^s\left(p_{u,-j} - p_{u,j} + \underline{\lambda}(\overline{p}_a - p_{a,j})\right) \qquad \text{(A9)}\\ &\quad + (1-\alpha)\left(p_{u,j} - c_u + \overline{\lambda}\left(p_{a,j} - c_a\right)\right) D^m\left(p_{u,-j} - p_{u,j}\right),\end{aligned}$$

Here, for the same reason as in the fully unshrouded equilibrium, the firm's optimal strategy is to charge $\overline{p}_a$ when sophisticated consumers expect its competition to charge $\overline{p}_a$, allowing it to lower its upfront price and shift out the myopic consumers' demand. Given that $p_a^* = \overline{p}_a$, the firm's optimal upfront price is unchanged.

In response to its competitors shrouding their add-on prices, the firm is thus indifferent between shrouding and unshrouding; shrouding yields the same profit as unshrouding. This result differs from the equilibrium derived in Gabaix and Laibson (2006) for the case of avoidable add-ons; here, since sophisticated consumers know that they cannot substitute away from the add-on for certain and expect the add-on's price to equal the walk-away price, there is no gain to shrouding. While the shrouded equilibrium would be the optimal

choice if there were even a small cost to unshrouding and releasing the add-on price to sophisticated consumers, our exposition in section 2 relies on the unshrouded equilibrium for simplicity.

B. Calculation of Marginal Costs

In this Appendix, we describe in detail how we construct our estimate of marginal costs for the schools' three services.

The marginal cost of the upfront service covers five cost components: the fees for taking the theory and practice exams (Fee_{theory} and $Fee_{practice}$), the fee for purchasing the instructional materials provided to the student ($BookFee$), the salaries of the instructors ($InstWage$), and the per-km cost of operating a fleet car ($VehOpCost$), scaled by the kilometers a student drives during the instructional courses. The theory repeat course generates as cost only the fee that the school pays to the exam center for the student to take the exam one more time; this fee does not vary with the number of times the student has already taken the exam. The school does not require students to either retake classes or study additional materials at a cost. In contrast, the practice repeat course involves additional driving lessons, which generate scaled-down salary and vehicle operating costs to the school, before the student retakes the exam, for which the school covers the exam fee. Accordingly, we specify the marginal costs for the three services as:

$$\begin{aligned} MC_{Base} &= Fee_{theory} + Fee_{practice} + BookFee + InstWage \\ &+ (700 + 2DistCtr)\, VehOpCost \\ MC_{TheoryAddon} &= Fee_{theory} \\ MC_{PracAddon} &= Fee_{practice} + \frac{6}{33} InstWage + \left(\frac{6}{33}700 + 2DistCtr\right) VehOpCost, \end{aligned} \tag{A10}$$

and estimate each cost component as:

1. Exam adminstration fees, Fee_{theory} and $Fee_{practice}$. The IMTT provided us with information on the fees each of the 25 exam centers charges to a school for administering the theory and practice exam. Administration fees range from €15 to €21 with a mean of €17.5 for the theory exam, and from €30 to €52 with a mean of €41.0 for the practice exam. Private exam centers charge slightly higher fees than publicly run exam centers. For a given school, we use the exam administration fees from the exam center used by the school in calculating marginal cost. Some schools use several exam centers, in which case we employ the highest of the centers' fees.

2. Instructional materials expenses, $BookFee$. Numerous sources quote €10 per student for instructional materials such as driver handbooks, CD-ROMs, and so on.

3. Instructor Wages, $InstWage$. The industry is subject to a price floor for instructor pay whereby the minimum pre-tax instructor salary is set at €683.05 per month. This floor is not binding; interviews with ANIECA and other driving school representatives suggest that salaries range from €750 to €950. Within this range, we assume that instructor salaries are proportional to average monthly earnings in the school's municipality, to control for cost-of-living differences. Schools also pay a 23.75 percent social security tax and a €3.4 per diem stipend, which we include to calculate the schools' monthly all-inclusive labor cost per instructor.

   Since each student's base and practice repeat courses include 32 and five driving lessons, respectively, we convert the monthly salary to an hourly basis. ANIECA's instructor wage cost template suggests that instructors work eight-hour days for on average 18.33 days per month. The marginal labor cost a school incurs for each student in the base driving course then consists of 33 hours times the hourly rate, covering the required practice lessons and an additional hour for the actual driving exam that the instructor participates in. Note that this assumes that schools incur only fixed, but no marginal, costs for the classroom time that the student spends preparing for the theory exam, given that schools rarely operate at capacity. For the average school, this results in all-inclusive marginal labor costs of €237.10 per student for the base course, with an interquartile range of €25.36 around the median of €237.60. For the practice repeat course, we simply scale this $InstWage$ by 0.18 since the instructor now only spends six hours, instead of 33, with the student in five driving lessons and a driving exam. The wage component of the practice add-on marginal cost is thus only €43.11, on average.

4. Vehicle Operating Costs, *Distance driven times* $VehOpCost$. The most sizable marginal cost component stems from the usage of the driving schools' fleet during driving lessons. According to ANIECA's cost template, vehicle operating costs are comprised of (i) gasoline costs, (ii) depreciation costs, (iii) maintenance and repairs costs, (iv) tolls and other road fees, (v) and other costs. From these five cost components we generate a cost per kilometer of operating a driving school vehicle, $VehOpCost$, which we then multiply by the 700 kilometers that school owners state a student covers during his base driving course in lessons and the driving exam. For the practice repeat course, we scale the distance driven by 0.18, as with wages, to 127.3km. In both instances, we then add two times the distance to the nearest exam center covered in driving

the student to and from the practice exam.[18] We calculate the cost per kilometer as follows:

- *Gasoline costs.* We obtained diesel gasoline price information from the Direcção-Geral de Energia e Geologia. The agency was able to provide us with the most recent, at the time of our request of March 12, 2012, prices of all gas stations in a school's municipality. We are thus able to capture cross-sectional, but not time-series variation in the cost of gasoline. As the cost per liter of diesel, we employ the average of the five lowest gas prices in the municipality, or €1.42 for the average driving school. School owners and ANIECA quoted a gas consumption of six to seven liters per 100 km for the smallest fleet car. We scaled this figure using school-specific information on the median weight and horsepower of cars in the driving fleet, relative to the sample in total, to reflect that school fleets with greater median weight and median horsepower require more gasoline per km. These two pieces of information allow us to estimate an average gasoline cost of €.10 per km, ranging from €.09 to €.11 per km across schools.

To estimate the non-gasoline vehicle operating cost components, we require an estimate of the total annual usage of a typical fleet car. Given IMTT's information on the sizes of each school's fleet and student body, as well as the distance traveled per student, we estimate that a vehicle at the average school travels 21,059 kilometers annually. School and fleet size introduce variation in this estimate, which ranges from 14,354 to 30,307 km per car per year. We use this annualized average total number of kilometers traveled as follows:

- *Depreciation costs.* In line with ANIECA suggestions, we assume that each car in the fleet costs €20,000 and depreciates fully by the end of eight years. We use the estimated annual kilometers driven, combined with an assumption of constant annual usage, to distribute the car's value over its total lifetime usage and arrive at a cost-per-kilometer figure. The average depreciation cost is €.12 per km, and ranges from a low of €.07 to a high of €.19 per km.
- *Maintenance and repair costs.* Again following ANIECA, we assume that each car annually requires €1,300 in routine maintenance costs and two major repairs. Using the information on the characteristics of each school's fleet relative to the average car in the sample, we adjust this figure to recognize that older, heavier,

[18] We allow for some car pooling as suggested by driving school owners whereby for every 20 kilometers in distance to the exam center, an additional student shares the ride to the exam center.

and more powerful cars require more expensive maintenance and repairs. We convert the resulting total maintenance expense to a maintenance cost per kilometer by dividing by the total kilometers traveled per year. The average maintenance and repair cost is €.06 per kilometer, and ranges from a low of €.04 to a high of €.09 per kilometer.

- *Tolls and other road fees.* According to ANIECA, each car accrues €625 in such fees annually, which we then scale by total kilometers traveled estimate to obtain a per kilometer cost figure. The average toll and other road fee cost is €.03 per kilometer, and ranges from a low of €.02 per kilometer to a high of €.04 per kilometer.
- *Miscellaneous costs.* ANIECA suggests that each car annually accrues an additional €2,750 in other expenses, which we again convert to a per-kilometer cost estimate. The average miscellaneous cost is €.13 per kilometer.

Taken together, the sum of these five cost components yields our vehicle operating costs per kilometer. When multiplied by our estimate of the distance covered by the average student during the base driving lessons and exam, we obtain a marginal cost due to total vehicle operating expenses of €327.2 per student for the average school, with lower and upper quartiles of €294.6 and €356.1. For the practice repeat-course, the vehicle operating expenses amount to €55.70, on average.

We use these inputs to calculate each school's marginal cost for its upfront and add-on services according to Equation A10. Table A-1 summarizes the inputs into the cost calculation and the resulting totals. The cost of serving an additional student to a school is €632.7 on average for the base course, but ranges from as low as €515.6 to as high as €761.7 per student.[19] Approximately the same amount of variation in marginal costs is explained by variation across schools within the same market and across markets, with the standard deviations amounting to 69.9 percent and 71.9 percent of the aggregate standard deviation for within and between variation. The cost of providing the add-on services is not surprisingly much less; the average school incurs €17.4 in marginal costs for the theory add-on and €139.77 for the practice repeat course.

Since a number of the cost inputs relied on the size and composition of each school's fleet, the information on which is updated by IMTT only periodically, we consider as a

[19] One driving school owner quoted a cost of €20 per student for his average driving lesson (which includes instructor wages and vehicle operating costs only). When scaled to the mandatory 32 lessons, this yields an estimated €640 per student, close to our estimated average marginal cost of €632.7 per student. Note however that our estimate also includes the instructional materials and exam fees.

Table A-1: Marginal Cost Components for Upfront Service (€)

			Distribution				
	Mean	Std Dev	Min	Q25	Med	Q75	Max
Upfront Service							
Exam administration fees	58.45	10.23	45	45	59	73	73
Instructional materials	10	10	10	10	10	10	10
Instructor wages	237.09	18.26	204.21	222.93	237.56	248.29	273.34
Vehicle operating costs	327.17	43.90	237.05	294.55	324.44	356.08	440.85
Total	632.70	48.91	515.55	594.38	630.31	669.37	761.71
Constant-usage MC, Total	631.27	17.79	579.88	619.24	632.16	642.67	679.79
Theory Add-On							
Theory exam fee	17.45	2.32	15	15	17	21	21
Total	17.45	2.32	15	15	17	21	21
Practice Add-On							
Practice exam fee	41.01	8.13	30	30	42	52	52
Instructor wages	43.11	3.32	37.13	40.53	43.19	45.14	49.70
Vehicle operating costs	55.66	8.80	36.99	49.26	55.02	61.76	81.07
Total	139.77	12.46	109.50	130.84	139.84	149.20	168.30
Constant-usage MC, Total	139.55	9.24	115.46	131.89	140.76	146.28	161.66

Note: The constant-usage MC is an alternative cost measure based on a constant annual distance covered per driving fleet car of 20,720 km in approximating each school's vehicle operating costs.

robustness check an alternative marginal cost measure that employs, for each school, the same estimate of the total annual distance traveled per fleet car. This also addresses the fact that depreciation expenses might reflect total usage, in addition to the age of the car, and that our current spreading of a constant depreciation amount over the car's total usage underestimates depreciation expenses for heavily used vehicles. Using the median total of 20,720 km traveled per year and fleet car for each school results in a marginal cost estimate for the upfront service of €631.27 on average, similar to our primary estimate. This measure of marginal cost, which we term constant-usage MC in Table A-1 below, has not surprisingly a lower standard deviation of €17.79 instead of €48.91. In absolute terms, however, the variation in marginal cost is low when compared to prices: our primary marginal cost estimate exhibits only half as much variation as that of the upfront price. As a result, variation in markups is primarily driven by variation in prices, rather than cost, and our final markup calculations are robust to using the constant-usage marginal cost.

C. Nonlinear Specification Estimation Procedure

In this Appendix, we describe how we construct and estimate the likelihood for our nonlinear specification. We estimate a system of two nonlinear equations using full information maximum likelihood (FIML), which predicts schools' chosen markups (resulting from their choice of prices) while controlling for correlations in unobserved market attributes that render market entry endogenous. Our estimation accounts for the fact that the markup is continuous while the market structure is discrete.

We specify the choice of markups for school i in market m as a (continuous) random effects model:

$$p_{im}^P = f^P\left(\alpha^P, \beta^P, \gamma^P, \mathbf{X}_{im}^P\right) + \xi_{im}^P \qquad \text{(Markup Equation)}$$

Here, we define $f^P(\cdot)$ as

$$f^P\left(\alpha^P, \beta^P, \gamma^P, \mathbf{X}_{im}^P\right) \equiv \alpha^P + \beta^P \mathbf{X}_{im}^P + \sum_{j=2}^{6} \gamma_{jm}^P \mathbb{1}_{\{\Pi_m^E = j\}},$$

where $\mathbf{X}_{im}^P$ contains school- and market-specific attributes. We estimate the market structure random effect, γ_{jm}^P, for markets with at least two firms in order that we have repeated observations by market.

Identifying the causal effect of entry on markups would be difficult since both markups and entry reflect the attractiveness of a market in difficult-to-measure ways. We control for the possibility of endogeneity by considering entry prior to deregulation in 1998. That is, we instrument for the number of entrants using the number of competitors prior to deregulation, which has relevance to the current number of entrants, but is otherwise uncorrelated with incumbents' current markup decisions.

We specify the number of schools (or entrants) in market m, which ranges from 1 to 6 in the data, as a (discrete) ordered probit model:[20]

$$\Pi_m^E = \begin{cases} 1 & \text{if } f^E\left(\alpha^E, \beta^E, \mathbf{Z}_m^E\right) + \varepsilon_m^E < \zeta_1^E \\ j & \text{if } \zeta_j^E < f^E\left(\alpha^E, \beta^E, \mathbf{Z}_m^E\right) + \varepsilon_m^E < \zeta_{j+1}^E \text{ for } j = 2, .., 5 \\ 6 & \text{if } f^E\left(\alpha^E, \beta^E, \mathbf{Z}_m^E\right) + \varepsilon_m^E > \zeta_6^E, \end{cases} \qquad \text{(Entry Equation)}$$

where the parameter ζ_j^E implies a cutoff for the unobservable ε_m^E between $j-1$ and j schools

[20] There are several markets in the data for which there are more than six schools, but these markets are few. We combine markets with six or more schools into the final cutoff value.

in the market and we define $f^E(\cdot)$ as

$$f^E\left(\alpha^E, \beta^E, \mathbf{Z}_m^E\right) \equiv \alpha^E + \beta^E \mathbf{Z}_m^E,$$

where $\mathbf{Z}_m^E$ contains market-specific attributes as well as the pre-deregulation number of schools, which we use as an instrument for the potential endogeneity.

We assume the markup choice error term ξ_{im}^P can be decomposed into a market- and school-specific component with $\xi_{im}^P = \varepsilon_m^P + \eta_{im}^P$, where ε_m^P is the market-specific error term for markups and η_{im}^P is the school-specific error term for markups with $\eta_{im}^P \sim N\left(0, \sigma_{\eta,P}^2\right)$. Normalizing the variance of the entrant equation ε_m^E to one, we allow for a flexible correlation structure between the market-specific errors. Specifically, ε_m^P and ε_m^E are distributed bivariate normal as follows:

$$\begin{pmatrix} \varepsilon_m^P \\ \varepsilon_m^E \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_P^2 & \sigma_{PE} \\ \sigma_{EP} & 1 \end{pmatrix}\right) \tag{A11}$$

Here, ε_m^P and ε_m^E are unobservable, market-specific factors that affect the entry and markup decisions, respectively, of all schools. The covariance terms allow for the kinds of correlations in the market-level unobservables discussed above.

We estimate three distributed-related parameters, $\left(\sigma_{\eta,P}^2, \sigma_P^2, \sigma_{EP}\right)$. In addition, we estimate the coefficients on the markup equation covariates, $\left(\alpha^P, \beta^P, \gamma^P\right)$, which includes the market structure indicator variables, γ^P, the coefficients on the entrant equation covariates, $\left(\alpha^E, \beta^E\right)$, and the ordered probit cutoff estimates, ζ_j^E.

In estimating the nonlinear system of equations with schools' markups as our dependent variable, the contribution of the likelihood from market m is the likelihood function

$$L_m = \Pr\left(p_{im}^P = p_i \forall i, \Pi_m^E = j\right),$$

where j is an index of the observed number of entrants and p_{im} equals the observed markup of school i. Thus, this likelihood of observing schools' markup choices in market m equals the joint probability distribution given by

$$\Pr\left(\xi_{im}^P = p_i - f_i^P, \zeta_j^E - f^E < \varepsilon_m^E < \zeta_{j+1}^E - f^E\right),$$

where ζ_j^E is the cutoff for j entrants. This probability is given by the integral of the $2N+1$-dimensional normal distribution of ξ_{im}^P and ε_m^E with mean zero and variance-covariance matrix

given by (where $\mathbf{I}$ is the identity matrix and $\mathbf{\Xi}$ is a matrix of ones)

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_P^2 \mathbf{\Xi}_{2N\times 2N} + \mathbf{I}_{2N\times 2N} & \sigma_{PE}\mathbf{I}_{2N\times 1} \\ \sigma_{EP}\mathbf{I}_{1\times 2N} & 1 \end{bmatrix}$$

over the surface defined by f^P; and f^E and the cutoffs ζ_2^E through ζ_6^E that are consistent with observed markups and the observed number of entrants, respectively. Note that $\mathbf{\Sigma}$ results from stacking the $2N$ markup decision errors $\xi_{im}^P = \varepsilon_m^P + \eta_{im}^P$ and the single market-level error, ε_m^E. The variance-covariance matrix allows for correlation in the unobservable market shifters of the markup and entry equations, and thus controls for the endogeneity of market structure across equations.

The assumption that $\xi_{im}^P = \varepsilon_m^P + \eta_{im}^P$ with $\eta_{im}^P \sim N\left(0, \sigma_{\eta,P}^2\right)$ allows us to simplify the likelihood by integrating out η_{im}^P, which results in the likelihood function being equal to

$$L_m = \int_{\zeta_j^E - f^E}^{\zeta_{j+1}^E - f^E} \int_{-\infty}^{\infty} \left[\prod_{i=1}^{N} g_i^P\left(\varepsilon_m^P\right)\right] \phi\left(\varepsilon_m^P, \varepsilon_m^E\right) d\varepsilon_m^P d\varepsilon_m^E,$$

where

$$g_i^P\left(\varepsilon_m^P\right) = \phi\left(p_i - f_i^P - \varepsilon_m^P\right)$$

is the standard normal pdf of η_{im}^P and $\phi\left(\varepsilon_m^P, \varepsilon_m^E\right)$ refers to the pdf of the bivariate normal distribution of $\left(\varepsilon_m^P, \varepsilon_m^E\right)$ given by A11.

We further integrate ε_m^E out of the likelihood, conditioning on the markup equation market error ε_m^P to obtain

$$L_m = \int_{-\infty}^{\infty} \left[\prod_{i=1}^{N} g_i^P\left(\varepsilon_m^P\right)\right] \times \left[\Phi_{\varepsilon_m^E|\varepsilon_m^P}\left(\zeta_{j+1}^E - f^E\right) - \Phi_{\varepsilon_m^E|\varepsilon_m^P}\left(\zeta_j^E - f^E\right)\right] \phi\left(\varepsilon_m^P\right) \varepsilon_m^P,$$

where $\Phi_{\varepsilon_m^E|\varepsilon_m^P}$ denotes the conditional cdf of ε_m^E, given realizations of ε_m^P.

For a given value of the parameters, we use simulation techniques to compute each market's contribution to the likelihood function by integrating numerically over the normal distribution of ε_m^P. This yields the continuous pdf of the markup distribution. We then use a numerical maximization routine to maximize the full likelihood equation

$$L \equiv \prod_{m=1}^{M} \left\{ \int_{-\infty}^{\infty} \left[\prod_{i=1}^{N} \frac{1}{\sigma_{\eta,P}} \phi\left(\frac{p_i - f_i^P - \varepsilon_m^P}{\sigma_{\eta,P}}\right)\right] \times \left[\Phi_{\varepsilon_m^E|\varepsilon_m^P}\left(\zeta_{j+1}^E - f^E\right) - \Phi_{\varepsilon_m^E|\varepsilon_m^P}\left(\zeta_j^E - f^E\right)\right] \phi\left(\varepsilon_m^P\right) d\varepsilon_m^P \right\}$$

and update the parameters until convergence.