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

Current version: October 2014

Katja Seim† Maria Ana Vitorino‡ David Muir§

Abstract

This paper empirically investigates the add-on or "drip" pricing behavior of firms in the Portuguese market for driving instruction. We present a model in which 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 loss leader pricing strategy emerges whereby firms price the upfront product below single-product pricing levels, but the add-on at monopoly levels. We then 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 evidence in support of the model predictions, including that firms have a profit motive in the add-on market. Prices for the base product, but not the add-on products, decline in the number of competitors a firm faces. Estimates of an empirical version of the model suggest that approximately one-quarter of students are inattentive to the add-on when making their school choice, closely matching the share of students who are uninformed about the add-ons' prices in a survey. This result has important policy implications about the cross-subsidization from those students who are unaware of the add-on to those who are.

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

JEL Classification: L10, L15, L80

*We thank Susana Paulino at Instituto da Mobilidade e dos Transportes Terrestres for access to the data and information about the industry. Ana Isabel Horta provided excellent research assistance, and Susana Belo assisted with local data collection. We also thank the participants of the 2012 Marketing Science conference (Boston), the Sixth Annual Federal Trade Commission Microeconomics Conference (Washington D.C.), the 2013 QME Conference, the 2013 UT Dallas FORMS conference, the Winter 2013 NBER Industrial Organization Meeting (Stanford), the 2014 IIOC conference (Chicago), and seminar participants at the Portuguese Competition Authority, Erasmus University (Rotterdam, Netherlands), Tilburg University (Netherlands), Stanford University, Penn State University, and Anderson School of Management (UCLA) for their comments. We specifically thank Glenn Ellison, Michael Grubb, Kanishka Misra, Stephan Seiler and Venkatesh Shankar for their invaluable comments. Maria Ana Vitorino gratefully acknowledges support from the Grant-in-Aid Program (GIA) at the University of Minnesota. All errors are our own.

†Wharton School (Business Economics & Public Policy Department), University of Pennsylvania, Philadelphia, PA 19104, kseim@wharton.upenn.edu.

‡Carlson School of Management (Marketing Department), University of Minnesota, Minneapolis, MN 55455, vitorino@umn.edu.

§Alfred Lerner College of Business and Economics, University of Delaware, Philadelphia, PA 19104, muir@udel.edu.

1 Introduction

In many industries, sellers advertise a low price for an 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" with search costs, price discrimination, or switching costs.[1] Gabaix and Laibson (2006) formalize another explanation for high markups in add-on markets: high add-on prices may be optimal when consumers do not anticipate that the price of the unadvertised add-on product is likely to be high. As an example, they point to tied products such as printers and toner cartridges in which consumer unawareness of the cartridge's high markup reduces the role that the tied product plays in the purchase process.[2]

To what extent are these theoretical predictions about add-on pricing borne out in reality? The empirical literature on add-on pricing is scant. Several recent papers' findings are consistent with products serving as loss leaders. For instance, Chevalier et al. (2003) demonstrate declines in retail price margins for certain grocery store products during peak demand periods, which suggests that these products serve as loss leaders in driving store traffic. 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 products at higher prices. In a field experiment, Chetty et al. (2009) find that demand falls when retailers "unshroud" by posting tax-inclusive prices for personal care products. In other field experiments, Hossain and Morgan (2006) and Brown et al. (2010) demonstrate that raising shipping charges increases revenues and the number of bidders attracted to eBay auctions.[3] There is little systematic evidence on the competitive interaction of firms in such environments, and the extent to which their behavior reflects optimal responses to consumer preferences, however.

In this article, we empirically investigate firm competition in the Portuguese market for driving instruction. Here students initially purchase a base course of driving instruction, completion of which entitles them to one attempt at passing a theory and a practical driving exam. Should the student fail either exam, additional fees accrue for both lessons and a repeat exam, the add-on service in this context. There are a number of reasons why students may be inattentive to repeat fees. First, while schools are required to keep a full schedule of their repeat fees on site, they are not required to inform students about their pricing structure at the time of registration. Second, as with professional testing and certification markets more generally, we show using survey evidence

[1] See Ellison (2005) for a general framework that intersects these three explanations.

[2] Gabaix and Laibson (2006) is related to a broader literature that explores how biases in consumers' beliefs explain various pricing practices. Particular features of subscription pricing are consistent with quasi-hyperbolic discounting (Oster and Scott Morton 2005; DellaVigna and Malmendier 2006), lack of self control (DellaVigna and Malmendier 2004) or overconfidence (Grubb 2009); loss aversion might motivate bait-and-switch pricing (Köszegi and Rabin 2006; Heidhues and Köszegi 2010); and boundedly-rational heuristics can be applied to study how market equilibria exploit noise in consumer product evaluations (Jin and Leslie 2003; Spiegler 2006).

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

that consumers systematically underestimate their probability of failing an exam and, thus, their demand for the add-on.[4] This is despite the fact that exam repetition is common; 44.2 percent of students fail either the theory or the on-road exam the first time.

A unique feature of our setting is that we observe the universe of Portuguese driving school students over a three-year period, including information on the school they attend, the sequence of their 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 school fees for the basic driving course and for repeating either the theory or the on-road portion. We also exploit the simplicity of the firms' cost structure, together with data on school-specific cost shifters, to estimate the per-student cost to the school for the basic and repeat courses and profit margins for schools' upfront and add-on services. Taken together, these data allow us to trace the entire sequence of purchases that students make, the prices they pay for each purchase, and the role of the competitive environment on pricing.

To study firms' pricing, we first set up a basic model that builds on the Gabaix and Laibson (2006) model of add-on pricing with inattentive 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 rational in forming expectations about their likelihood of needing to purchase the add-on, and can engage in costly effort to reduce their purchase incidence. Inattentive consumers, conversely, do not account for the add-on and their likelihood of requiring it when choosing their school. We show that a typical loss-leader pricing strategy emerges in which firms sell the upfront product below competitive profit margins and simultaneously price the add-on at monopoly 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 upfront losses. This feature has two consequences for the upfront product's price. First, a larger share of inattentive consumers, who unexpectedly (to them) participate and generate profit in the add-on market, depresses the upfront product's price. Second, an increase in the probability with which consumers require the add-on product similarly raises profitability, offset by a lower upfront product price and markup.

We then test several of these predictions in the driving school setting. First, we establish that the rates of failure across schools result in significant revenue and profit to schools from the add-on market; our data indicate that an estimated 16.8 percent of revenue and 31.6 percent of variable profit derive from repeat courses for the median school.

Second, while the percent markup for the base course averages 27.2 percent, the corresponding figures average to 86.4 and 57.5 percent in the theory and on-road add-on markets, respectively. Schools' upfront prices and markups strongly correlate with the number of schools in their municipality, pointing to the standard downward pressure on prices of additional competition. On the

[4] In consumer financial markets, consumers might similarly underestimate their future need for account features such as overdraft services or financing, at the time when they open a bank or credit card account (Stango and Zinman 2009).

other hand, schools' repeat fees and markups do not correlate significantly with market structure, as predicted by the theoretical model.

Since our data do not contain direct evidence on students' ex-ante perception of either their likelihood of failing or their understanding of the financial repercussions of having to retake the exam, we complement our primary data with survey evidence from a subsample of students who completed a questionnaire as part of their theory exam session at a select testing center. The survey evidence suggests that students overestimate their probabilities of passing a yet-to-be-taken exam. At the same time, the survey responses identify a sizable group of students – approximately 22 percent – who either do not know whether they will or believe that they will not incur any fees for retaking an exam. These are two possible reasons why students would discount the price of the add-on at the time of their school choice.

In addition to providing descriptive evidence in support of the model's predictions, we estimate an empirical version of the model that relies on students' responses to schools' fee structures to back out the proportion of students whose school choice is consistent with a failure to account for the add-on prices. We detail a two-type mixture model of school choice in which the two types, as in the theoretical model, differ in their anticipated probabilities of failing an exam. They make school choices based on whether they consider the upfront price only or the full price, which includes the expected repeat fees. Our results demonstrate that schools' pricing behavior does not merely reflect students' differing price sensitivities across product markets; rather, schools' pricing is consistent with approximately one-quarter of the student population disregarding repeat fees when making their school choices.

That inattentive students constitute a nontrivial subset of the student population has important distributional consequences. If sophisticated students can reduce their demand in the add-on market, they reap the benefits of low prices in the upfront market. Thus, inattentive students – whose business in both markets comprises a significant portion of schools' profits – effectively subsidize the remaining students with lower prices in the upfront market.

The paper proceeds as follows. Section 2 presents a basic loss-leader model with inattentive students. Section 3 introduces the data and the institutional setting of Portuguese driving schools. Section 4 and section 5 provide evidence to support some of the basic predictions and implications of the model based on our observational and survey data. Section 6 formalizes an empirical model using several of the model implications. Section 7 concludes.

2 A Model of Add-on Pricing with Inattentive Students

Here, we outline a simple model of add-on pricing in the spirit of Gabaix and Laibson (2006) and Spiegler (2011) to illustrate that the loss-leader pricing strategies common to multi-product settings with search or switching costs can also arise when consumers are inattentive to 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. 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. 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. As in classic repeat purchase models of pricing with switching costs (Klemperer 1987a; Beggs and Klemperer 1992; Farrell and Klemperer 2007), we assume that consumers are locked into purchasing both the upfront course and a possible repeat course from the same driving school.[5] In contrast to these models where firms are not able to commit to prices for subsequently purchased add-on products in the initial period, firms are required to keep at hand a full schedule of prices. We thus assume that schools commit to the add-on price when setting the price menu in the initial stage of the game.

In period 1, each school j simultaneously chooses and commits to a pricing strategy (p_j^u, p_j^a), where p_j^u and p_j^a 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, he might choose not to continue with driving instruction or lodge a complaint with the regulatory body, the IMTT. While the IMTT does not directly regulate the price of repeat courses, its oversight likely also limits the fees that schools can charge. In period 2, students learn whether or not they need to buy the add-on at its set price depending on their exam results from period 1. 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 $\pi \in (0, 1)$ sophisticated types s and $(1 - \pi)$ inattentive types m. In period one, the inattentive type disregards the add-on service; they only become aware of it ex-post when they fail the exam and are forced to purchase it. This assumption might reflect that inattentive students assign zero probability to failing the exam – the students suffer from over-optimism, for example – or that students are unaware of repeat 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 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. Sophisticated students can engage in costly effort to reduce their probability of failing from $\overline{\lambda}$ to $\underline{\lambda} > 0$ at an effort cost e.[6]

In line with our empirical setting, we follow Gabaix and Laibson (2006) in assuming that students make a discrete school choice, allowing for heterogeneous valuations of each school. Here, we assume that there are no systematic differences in valuations by type; we relax this assumption

[5] We discuss below that we observe a negligible share of school transfer by students in our data, which is consistent with the lock-in assumed here.

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

in Section 6. The utility of student i of type $\{m, s\}$ from enrolling at school j is given by

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

where ε_{ij} denotes student i's heterogeneous valuation for the school j, such as the distance he travels to the school, and $\lambda = \{\underline{\lambda}, \overline{\lambda}\}$, depending on whether the sophisticated student chooses to engage in effort.[7]

Individual 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 type I extreme value, this assumption results in common multinomial logit school choice probabilities:

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

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

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

Now consider pricing of the base service. In the Appendix we show that there is a unique symmetric equilibrium characterized by the following pricing strategies:

$$\begin{aligned} (p_j^a)^* &= \overline{p}^a \\ (p_j^u)^* &= c^u + \frac{\sigma n}{n-1} - \left[(1-\pi)\overline{\lambda} + \pi\underline{\lambda}\right](\overline{p}^a - c^a) \end{aligned} \tag{3}$$

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 charge identical add-on service prices equal to the monopoly price.

2. The add-on price does not depend on the fraction of inattentive types or the probability of failing.

[7] 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 sophisticated students form rational expectations about add-on prices and unshrouding is costless. We therefore simply assume that sophisticated students 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.

3. The price of the upfront service is decreasing in the fraction of inattentive types, or

$$\frac{\partial (p_j^u)^*}{\partial \pi} = \left(\overline{\lambda} - \underline{\lambda}\right)\left(\overline{p}^a - c^a\right) \geq 0. \tag{4}$$

Thus, as the size of the inattentive segment increases, the upfront price decreases to reflect that schools anticipate larger profits from the add-on service; schools want to attract more inattentive types upfront and recoup these losses in the add-on market.[8]

4. The price of the upfront service is monotonically decreasing in the average probability of failing the exam, $\left[(1-\pi)\overline{\lambda} + \pi\underline{\lambda}\right]$. The greater the number of students schools can attract in the add-on market, the lower the price they charge in the upfront market to entice students to sign up with with them.

5. Prices in the upfront market decline as the number of firms increases. With large n and a sizable add-on market, 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 (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 $\left(\overline{p}^a - c^a\right)\left(\overline{\lambda} - \underline{\lambda}\right)$ on every sophisticate. The choice to engage in effort, thus, is only efficient if $e \leq c^a\left(\overline{\lambda} - \underline{\lambda}\right)$, whereas the student's choice to do so reflects the prices he 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, any individual aged 18 years or older must first enroll in an IMTT-authorized driving school.[9] There, candidates must complete 28 theory lessons, the curriculum of which is set by the IMTT, and a minimum of 32 on-road driving lessons. After completing the required theory lessons, students take a computerized theory exam. Subsequently, they perform an

[8]This conclusion depends on the relative ordering of the probabilities of failing by inattentive and sophisticated types. If sophisticated students had a strictly higher probability of failing than inattentive ones, regardless of their expended effort, the price of the upfront service would increase in the inattentive segment share.

[9]Licensing by the IMTT requires, among other things, proof that the proposed school owner holds 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.

on-road driving test. If a candidate fails either the theory or the on-road exam, he must pay the school a fee to retake the exam and, in the case of the on-road portion, to complete five additional driving lessons. Both exams are administered at one of 35 exam centers. Twenty-two exam centers are managed by the IMTT, while private organizations operate the remainder. An IMTT certified examiner oversees the on-road driving test.

As of the end of 2011, there are 1,141 driving schools in mainland Portugal. Since 1998, the industry has more than doubled relative to its initial size of 508 schools. Entry resulted from significant liberalization efforts that lifted restrictions on the number of schools serving each municipality to be tied to population: each licensed school had to serve a minimum of 25,000 residents or was the sole provider in a municipality with less than 25,000 residents. We exploit the previously regulated market structure as an instrument for today's market structure in some of the empirical analyses below.

Firms typically charge a single fee for the base course of instruction, covering classroom time, materials, practice theory exams, on-road driving lessons, and exam administration by the exam center, and separate fees to retake any portion of the course. 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. Eighty-seven percent of owners operate a single school, and another nine percent operate two schools. Figure 1 plots the locations of all driving schools in mainland Portugal by municipality population density.

========================

Insert Figure 1 about here

========================

3.2 Data

Our empirical analysis combines a number of data sets at the school, student, and municipality level. First, from the IMTT, we obtained data on school characteristics related to its instructors and driving fleet. These data originate from the school's licensing application and are updated periodically. We geocoded school addresses using GIS software and added hand-collected school prices and estimated costs. Second, IMTT provided individual-level school enrollment and driving exam information for the universe of students who obtained their driver license's in 2009 or 2010. Finally, we obtain demographic and other data at the level of the parish or municipality from the market research company, GrupoMarktest, the Ministério do Trabalho e da Solidariedade Social, and Statistics Portugal.

Our empirical analysis treats a municipality as the relevant market area within which firms compete and students choose schools. In line with this assumption, we exclude the districts of

Lisbon and Porto from the sample. While schools are required to conduct their business within the municipality covered by their license, students can attend a school outside of their municipality of residence. Since Lisbon and Porto are large, densely populated districts, the assumption of a single municipality comprising the relevant market area is less reasonable there. We further restrict our sample to schools in municipalities where complete price information is available for all schools, as discussed below.

On the student side, the original IMTT records contain applicants for all license categories. We focus on applicants for a type-B passenger-vehicle license, resulting in a loss of 16.3 percent of candidates. Further, we restrict the sample to candidates for whom we have a complete history profile from enrollment at a school to completion of the on-road exam, which eliminates another 25.8 percent of the full data set due to left or right truncation. Eliminating students residing in Lisbon or Porto reduces the number of applicants by another 25.8%.

The final sample contains 90,446 students residing in one of 214 municipalities with a total of 636 schools. These are displayed in panel (b) of Figure 1.

3.2.1 Sample Markets

Table 1 describes the sample of markets. The average (median) municipality contains 3.2 (two) driving schools, ranging from one to 22 schools. They range in population from 2,952 to 181,474 residents, with an average population of 27,158 people and have 422.6 driving school students during our data period. The average municipality is broken down into 24.8 parishes. 80% of residents live in a parish that is either moderately or predominantly urban. Monthly wages amount to €861.25. 12.8% of residents have completed some level of post-secondary education. The average monthly rent of €861.25, the availability of commercial properties proxied by the share of non-residential buildings of 8.9%, and the number of auto repair shops of 62 on average serve as shifters of the schools' costs.

=========================

Insert Table 1 about here

=========================

3.2.2 School Characteristics

Here, we summarize school characteristics, the size of the schools' student body, and market shares. See Table 2. For the average school, 142.2 students completed their course during our sample period.[10] There is significant variation in enrollment figures across schools, however; the interquartile range of enrollment spans from 87 to 176 students.

[10] Note that the student body of the typical school is significantly larger due to the fact that we restrict ourselves to the sample of students who began and completed their course in our sample period. In total, over the three year period, the average school enrolled 425.9 students, or approximately 140 students per year.

==========================
Insert Table 2 about here
==========================

The schools employ 5.8 instructors who have worked in the firm for a total of eight years on average. The schools have a driving fleet consisting of, at the median, three passenger vehicles, for which we observe characteristics such as displacement, age, and weight. We calculate the straight-line distance from the school to the nearest district-wide IMTT office and to its most used exam center to proxy for costs of interacting with the IMTT and of transporting students to the exam center. The average school is 23.4 kilometers from the closest IMTT office. It also is 21.1 kilometers from its most commonly used exam center. At the average exam center, eight examiners administer on average 57.9 exams per month.

The IMTT does not collect price information. Instead, we complement the IMTT data with hand-collected, detailed prices. We employed a team of 14 mystery shoppers who visited each school in person between November 2011 and March 2012 with an identical script to obtain information on base prices and repeat fees.[11]

The bottom panel of Table 2 summarizes the distribution of prices across schools and markets. The median school charges €700 for its base driving course, with an interquartile range of €150. 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 2.2 times the within-municipality standard deviation in upfront prices. The average on-road repeat course fee of €275.2 is more than double that of the theory repeat course fee, €129.1, which is driven by the fact that the on-road repeat course includes five additional lessons.

3.2.3 Marginal Costs

The stylized model above makes predictions for pricing patterns for base and repeat courses under the assumption of constant marginal cost. To test whether any patterns in prices are driven by variation in cost, we estimate schools' marginal costs. We benefit from the simplicity of the service offered and exploit information contained in a template that the industry association Associação Nacional dos Industriais do Ensino da Condução Automóvel (ANIECA) provides to members to estimate annual operating costs, including both total and unit costs.

According to the cost template, the base course marginal cost per student consists of (i) fees paid to the exam center for one theory and one on-road 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 (which he needs to attend), and (iv) the vehicle operating costs of driving one of the school's vehicles during the practice lessons (cost of gasoline, depreciation expenses, maintenance and repairs, tolls and other road fees and taxes, and other expenses). The theory

[11] Out of the 245 municipalities outside of the Lisbon and Porto districts, we dropped 31 municipalities due to incomplete price information for at least one school in the municipality

repeat course only entails as cost to the school the student's theory exam fee, which is identical to the fee it pays to the exam center the first time the student takes the exam. For the on-road repeat course, the school incurs exam administration fees, as well as labor and vehicle maintenance costs equivalent to five driving lessons.

We use information on exam center fees, municipality-level salaries, local gasoline prices, the estimated distance traveled in kilometers during a 32-lesson and a five-lesson course of instruction, and the annual usage in kilometers of each school's fleet of cars to derive marginal cost estimates (see Appendix B for details). Per student, the average school pays €53.02 in exam fees to cover the cost of one theory and one on-road exam. The cost of instructional materials is minimal and standardized, amounting to €10 per student. We estimate that the school's labor cost of an instructor amounts to €235.6 and €42.8 for 32 and five on-road lessons, respectively; we assume no marginal labor costs arise from the instructor's classroom time due to the common spare capacity in classroom space. The variation in labor costs captures cost-of-living differences reflected in municipality-level incomes across municipalities. Vehicle operating costs represent the largest source of marginal costs; we estimate the costs of gasoline, depreciation, and maintenance and repairs to amount to €0.28 per kilometer.[12] When scaled by the 722.9 kilometers the average student covers during the driving course and exam itself, the vehicle operating cost for the base and on-road repeat courses amount to €194.9 and €32.9 on average.

In total, the average school incurs a marginal cost of €493.6 per student in the base course, with a standard deviation of €18.69. The add-on services generate estimated marginal costs of €16.37 and €112.42 for the theory and on-road repeat courses, respectively. See Table B.1. We verified the reliability of our cost estimates in interviews with driving school owners and using feasibility studies that potential entrants prepare for the IMTT as part of the licensing process. Our cost estimates are comparable to the new entrants' own cost estimates for a country-wide sample of schools.

3.2.4 Student Attributes

The IMTT student records contain the date on which the student obtained 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 the seven-digit postal code of his residence, which approximately designates a city block.

The average (median) student in our sample is 21.9 years (19.1 years) old. There is an even split of male and female students: 51.6 percent of our sample are females. We assume that each student resides at the centroid of his postal code area to compute his distance to his chosen school, and to the other schools in his consideration set.

We calculate the straight-line distance from the postal code to the school that the student has chosen to attend. According to this measure, not only are a majority of students located less

[12] This compares to estimates of vehicle operating and ownership costs provided by the Automóvel Club de Portugal (ACP).

than three kilometers to their schools (2.9 kilometers), but 46.4 percent of students choose the school closest to home. See Table 3. Clearly, in addition to competition from other schools, spatial differentiation is an important dimension of the observed price variation.

==========================
Insert Table 3 about here
==========================

4 Pricing in the Upfront and Add-on Markets

In this section, we investigate the extent to which prices conform to the predictions of the stylized model above, and describe the economic relevance of the retake markets more generally. We begin by summarizing students' exam outcomes, school pass rates, and thus the size of the repeat market.

4.1 Incidence of Exam Repetitions

Figure 2 and Table 3 suggest that the add-on market is sizable. 76.0 and 73.2 percent pass the theory and on-road exams on the first attempt, respectively. The two exams' outcomes are largely independent since only 55.8 percent pass both exams on the first attempt. The average student takes the theory and on-road exams approximately 1.37 and 1.36 times, respectively. This results in the driving school process being lengthy; the median student takes seven months from start to finish. From the school's perspective, the retake market has the potential to serve as a significant source of profit.

==========================
Insert Figure 2 about here
==========================

The data also provide evidence to support our model assumption that students are locked into the school by their initial choice and rarely switch schools. Only 1.8 and 0.1 percent of all students transfer schools during the theory and on-road exam phases, respectively, while only 0.7 percent transfer in between the theory and on-road exam phases. The majority of these, or 64.5 percent, transfer to schools outside their original school's municipality, suggesting that exogenous reasons such as moving explain a significant share of transfers. The primary switching cost is that lessons taken at one school do not transfer to another school if a student switches before obtaining his license; switching schools requires restarting the base course from the beginning. This is costly: given each student's current school choice, the cheapest base course of instruction at a different school in the municipality has a price equal to 5.4 times the current school's theory repeat fee and 2.4 times its on-road repeat fee, on average.[13]

[13]Given the low empirical incidence of school switching, we ignore the possibility of inducing it through primary good pricing. See Miao (2010) for a theoretical analysis of such simultaneous add-on and primary good competition.

The theoretical model also proposes that schools' add-on prices would be set at their market's "walkway" price, or the maximum supportable price in the market. The data provide evidence that this assumption is reasonable. First, conditional on failing at least one exam, students are more likely to quit than to transfer. While 0.9 percent of students at schools in our sample switch schools within the same municipality, more than double – or 2.6 percent of students – quit.[14] Second, conditional on failing at least one exam, students who quit also behave in ways we would expect. Conditional logit regressions of the decision to quit reveal that students are more likely to quit schools with higher repeat fees for the particular exam that they failed, a statistically and economically significant finding that is robust to adding school- and municipality-level controls. Note that, if price is correlated with other components of school quality not captured by our controls, the estimated price effect would be downward biased.

4.2 Markups

To assess the profitability of offering the three services, we consider estimated price-cost margins and the percentage markup or the ratio of markup to price. Table 4 summarizes both; Figure 3 illustrates the distribution of the percentage markups.

==========================

Insert Figure 3 about here

==========================

==========================

Insert Table 4 about here

==========================

The average percentage markup for the upfront service is 27.2 percent, while in contrast, the percentage markups in the add-on markets average to 86.4 percent and 57.5 percent in the theory and on-road add-on markets, respectively. The fact that markups are higher in the repeat than in the base markets is corroborated by the fact that in total, repeat fees amount to 58.5% of base prices for the average school, despite the fact that the theory retake requires no and the on-road retake only five (or 17.9% of the base course's requirement) additional lessons.

Ten schools, or 1.6 percent of the sample, have negative markups in the upfront market, pricing below estimated cost. All but two of these schools earn positive total profit 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.[15] We compute effective markups

[14] Quitters are defined as those students for whom more than a year passes between successive exams, conditional on failing the last exam taken, or as those students who disappear altogether after their last instance of failing an exam.

[15] We estimate that the two schools have negative markups of −€7.7 on average across the three services. We are unable to determine whether this is because of measurement error in cost or because this school truly incurs a loss.

as the total fees paid by each student less the total variable cost incurred in serving the student, averaging across students at each school. A school earns €313.8, or 33.8 percent of fees paid, on the average student.

The equilibrium of the stylized model suggests that the price determinants of upfront and add-on markets differ: the upfront market prices depend on a number of consumer- and market-related factors, while the add-on fees are independent of these. Empirically, there is only a weak correlation of the base course and the theory repeat course prices (0.071), or the base course and the on-road repeat course prices (0.083). There is, however, a stronger correlation between each of the add-on prices (0.352). These correlations provide initial evidence that factors affecting the setting of schools' upfront prices differ from those affecting their add-on prices.

4.3 Sources of Profit

We now translate the markups and exam repeat incidences described in the previous sections into profitability breakdowns. A significant portion of schools' variable profits derive from add-on fees. The median school derives 31.6 percent of profit overall from repeats broken down into 11.1 and 20.5 percent from the theory and on-road exam add-on services, respectively. At the same time, schools earn a smaller share of total revenues from add-on services: the median derives 6.2 and 10.6 percent of revenue from the two add-on services for a total of 16.8 percent of revenue. Thus, more revenue is passed on as variable profits to schools in the add-on market.

4.4 Determinants of Prices

The previous section establishes that schools have a profit motive in the add-on market in which significantly higher markups are earned relative to the base course. We now test whether the observed prices for the upfront and retake services are consistent with the basic predictions of the ex-post monopoly model from Section 2. The model predicts that prices for the base course of instruction, but not the repeat fees, vary with the number of competitors in the market, the proportion of inattentive types, and the types' probability of failing the exams. 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 clear proxies for the proportion of inattentive students. In Section 6 we estimate an empirical version of the demand model above, using the consumers' responses to variation in upfront and add-on prices to pin down the share of inattentive 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.

Figure 4 first summarizes the raw price patterns in the data showing a box plot of mean municipality upfront and repeat fees by number of competitors in the municipality. The plot is supportive of the first model prediction: upfront prices appear to decline in the number of competitors, while add-on prices do not. Since the plot does not control for other potentially confounding school and municipality characteristics, we test these predictions more formally in regression models of prices on market structure and the students' probability of failing, controlling for heterogeneity in demand and cost conditions across schools and municipalities. We employ both OLS and instrumental variables techniques to control for the possibility that the number of competitors is endogenous to school prices; for example, both entry and prices might be higher in markets with unobservably high demand, biasing the estimated coefficient on market structure upward.

We control for such potential endogeneity by employing the regulated number of schools per municipality prior to liberalization as an instrument: by being tied to population, it is likely uncorrelated with unobserved profit shifters that are reflected in prices in the unregulated regime, but is related to the number of schools in the municipality under free entry. We complement this primary instrument with fixed cost shifters of the school, including its distance to the local IMTT office, capturing the cost of doing business with the IMTT, as well as the proportion of nonresidential buildings and the average rent to capture office rental costs.

=========================

Insert Figure 4 about here

=========================

The results of the regressions of the upfront prices on the number of market competitors with school- and municipality-level controls are shown in Table 5. The first-stage F-statistics and partial R-squared statistics for the instrumental variables regressions imply a strong instrument at conventional levels, and the 2SLS results in column (3) enhance the negative effect of competition, relative to the OLS results, as expected. For every school added to the market, the upfront price decreases by approximately €10.87 (€9.52 under the OLS specification), corresponding to 2 percent of the average upfront price.

=========================

Insert Table 5 about here

=========================

Relaxing the linearity assumption of competition on price underlying specification (1) by including market structure dummies for each observed firm configuration suggests that the decline in prices is more pronounced in going from monopoly to duopoly markets, when prices drop by €28.2 in the OLS specification. The largest markets with six or more competitors have prices that are on average €85.1 lower than monopoly markets. To address endogeneity concerns in the nonlinear specification, we account for the fact that the endogenous variables are discrete market structure outcomes by modeling them as an ordered probit model, allowing entry to shift with municipality characteristics that are shared by the price model as well as the excluded instruments. To capture the source of endogeneity concerns, we introduce an unobserved market-level error in the price equation that is correlated with the error that drives entry decisions. We estimate the resulting two-equation model using full information maximum likelihood (FIML) and derive the likelihood in the Appendix. While not directly comparable to the OLS results (which did not include a market-level random effect in the pricing equation), the estimated price effects are similar in the instrumental variables results.

Across specifications, prices also respond to the share of students who fail both exams, our proxy for the add-on market size. A one standard deviation increase in the share of students who fail both exams (3.9 percent) is associated with a decline in price of between €6.56 to €10.04 across models. While this is consistent with the model prediction of shifting profit from the upfront to the add-on market, the effect is less pronounced than the effect of competition.

The results for the equivalent regressions of add-on prices on the number of market competitors with school- and municipality-level controls are shown in Table 6. We collapse repeat fees into a single sum of the two, which we employ as our dependent variable. As with the upfront price regressions, the results are robust to instrumenting for the number of firms and introducing school- and municipality-level controls. Across specifications, the number of schools in the school municipality is statistically insignificant in affecting the repeat fee, and has an economically smaller effect on price. Similarly, the size of the add-on market - the share of students failing both exams - does not have a significant association with repeat fess.

==========================

Insert Table 6 about here

==========================

As a robustness check, we repeat the analyses in Tables 5 and 6 using markups as the dependent variables to investigate whether estimated cost differences across schools explain some of the patterns above. The results in Appendix Tables D.1 and D.2 show economically similar and statistically stronger evidence that base markups, but not repeat course markups, decline in the number of competitors and the share of exam repeaters.

5 Survey of Students' Exam Outcome Expectations

The evidence presented so far supports the hypothesis that firms use loss-leader pricing in the upfront market, with significantly higher, and largely competition-invariant, margins in the add-on market. Since the data do not contain direct information on students' expectations of the cost or incidence of exam failures, we complement this analysis with a short survey administered to students in the process of obtaining their drivers licenses in the Setúbal district. We developed the survey jointly with the IMTT, and IMTT representatives administered the survey to all students who took the theory driving exam at Setúbal's public exam center during the months of December 2012 and January 2013. The students were given the opportunity to answer the survey immediately after having taken the theory exam, and prior to learning their outcomes on the exam. The survey was voluntary, and was presented to students as being part of a general study on driver education and testing. The IMTT shared with us the anonymized compiled survey responses, together with information on the students' ultimate performance on the immediately following theory exam and the later on-road exam. Figure 5 depicts the survey.

In total, 797 students took the theory exam in Setúbal during the two-months window of the survey, and 782 students participated in the survey, entailing a 98 percent response rate. In the following analyses, we focus on first-time theory exam takers of the passenger vehicle license exam only, resulting in a sample of 440 respondents. The resulting sample is representative of the main sample: 51.8 percent are female; the mean (median) age at the time of the theory exam is 22.6 (19.7) years with a standard deviation of 7.1 years. We use the survey responses to investigate two questions. First, do students understand the financial repercussions of failing the theory exam? Second, do students have correct expectations of their likelihood of failing the exam?

To speak to the first question, we asked the students whether they would have to pay an additional amount to retake the theory exam in case they failed the exam they just took. 4.1 percent of students stated they would not have to pay anything; 17.5 percent did not know; and 78.4 percent said that they would have to pay some amount. 26.4 percent of these students state that they asked the school whether they would have to pay, and 43.3 percent state that the school informed them directly that additional fees would accrue. However, only 66.4 percent of the students who expect to pay for a retake, or 52.0 percent in aggregate, stated that they knew the amount they would have to overpay. This amount if higher among female students, 56.6 percent of whom state knowledge of the amount of retake fees. See Table 7. The students who claim knowledge of the retake fees are close to correct, underestimating the retake fees by only 4.1 percent on average.

========================

Insert Table 7 about here

========================

Similar to the nationwide results, in the sample of survey respondents failing a driving exam is common: only 72.5 percent pass the theory exam at first try, while 76.6 percent pass the subsequent on-road exam at first try. In aggregate, students have near correct expectations regarding the likelihood of passing the theory exam at 69.1%: there are no statistically significant differences between the expected and the actual pass rate, in aggregate, by gender and by whether students are informed about the existence of repeat fees (categories 3 and 4 in Table 7) or not. One exception are female students whose expectation falls short of the actual exam outcomes by 8 percentage points. There is evidence, however, of optimism in passing the on-road exam. 86.6 percent of students state that believe they will pass the exam, compared to the 76.6 percent who ultimately do. Similar differences between expectations and outcomes persist in subgroups, with male and uninformed students having the largest differences. Note that students have different amounts of information when answering the two questions: their assessment of passing the theory exam is based on having just taken the exam, while for the average student in the sample, the driving exam takes place only 3 months after having answered the survey. To the extent that the students' initial school choice is similarly made without direct evidence of the difficulty of each of the exams, this suggests that students may underestimate their propensity of needing to purchase the add-on repeat course.

==========================

Insert Table 8 about here

==========================

Lastly, we use the survey to gain insight into the students' preparation for the exam, asking them on number and consistency of practice exams taken, use of the textbook, additional theory lessons taken, preparation of the the lessons' material before class, note taking, and interactions with the instructor (see Question 5 in 5). We use factor analysis to aggregate the student responses into two indices of preparation that we denote as "active" (loading heavily onto the question of interaction with instructors) and "passive" (loading primarily onto questions of use of book and note taking during class) preparation, as well as the share of the seven preparation categories that the student stated to have engaged in, which we term "overall preparation" in Table 9. Lastly, Table 9 reports students assessment of whether they needed to study only little in preparation for the theory exam. As Table 9 suggests, there are no significant differences between informed and uninformed students in the degree of active preparation. However, both informed and female students engage in higher levels of passive preparation; female students further prepare more overall and are less likely to feel that the test requires little preparation. To the extent that female students are also less likely to overestimate their probability of passing and are thus sophisticated, similar to informed students about test fees, these results are thus suggestive that sophisticated students engage in more effort in their exam preparation.

Primarily, however, the survey responses indicate that only a subset of students is aware of the fees they will have to pay if they need to repeat an exam and that students on average overestimate the likelihood of passing the on-road exam. These are two reasons for inattention to repeat fees in school choices.

6 Empirical Model

In this section we present and estimate an empirical model based on the theoretical model discussed in Section 2.

6.1 Setup

Each student makes a choice among all schools in his municipality based on a utility comparison. Student i's utility from attending school j is conditional on the student's type h (i.e., inattentive or sophisticated) and is defined as

$$u_{ij|h} = \mathbf{Z}_i \mathbf{X}_j' \beta_Z + f(D_{ij}) + \xi_j - \alpha \mathbb{E} p_{ij|h} + \varepsilon_{ij} = \hat{u}_{ij} - \alpha \mathbb{E} p_{ij|h} + \varepsilon_{ij}, \quad (5)$$

where $\mathbf{X}_j$ is a vector of non-price school attributes and $\mathbf{Z}_i$ is a vector of student-specific attributes such as demographics; D_{ij} is the distance from student i's location to the location of school j; $\mathbb{E}p_{ih|j}$ is the price that each student expects to pay for his training at each school; ξ_j denotes school fixed effects for all schools that belong to student i's choice set; and ε_{ij} is an unobservable (to the econometrician) error term that we assume to be mean independent of the included right-hand side variables.

To control for students' preferences for distance, we allow distance to enter nonlinearly into the utility function in a flexible way as

$$f\left(D_{ij}\right) = \beta_{d_1} D_{ij} + \beta_{d_2} D_{ij}^2 + \beta_{d_3} \mathbb{1}_{\{j=\text{closest}\}}, \tag{6}$$

where $\mathbb{1}_{\{j=\text{closest}\}}$ is an indicator variable equal to 1 if school j is closest to the student in the choice set of student i.

Since we do not observe individuals who chose not to obtain a driver's license in our data, we model students' choices between schools only and normalize one fixed effect in each municipality to zero.

The expected price $\mathbb{E}p_{ih|j}$ of attending school j varies by student i and is defined as

$$\mathbb{E}p_{ij|h} = p_j^u + \mathbb{1}_{\{h=s\}} \left(\hat{\pi}_{ij}^T p_j^{a,T} + \hat{\pi}_{ij}^P p_j^{a,P} \right) \tag{7}$$

where p_j^u is the price of the base course of instruction, $p_j^{a,T}$ the theory repeat fee, and $p_j^{a,P}$ the on-road repeat fee, all at school j. $\hat{\pi}_{ij}^E$ is student i's expected fail probability at school j for each of the two exam types $E = P, T$ for the on-road and theory exams. Student i considers expected fail probabilities only when he is sophisticated, indicated by $\mathbb{1}_{\{h=s\}}$.

6.2 Fail Probabilities

A key component to school choice involves the assessment that each sophisticated student makes of his probability of failing each exam at each of the schools in his choice set. We assume each student can fail each of the theory and on-road exams at most one time. This is consistent with the observation that most students rarely fail more than once. We estimate a conditional logistic regression model of the fail probability given by

$$\ln \left(\frac{\hat{\pi}_{ij}^E}{1 - \hat{\pi}_{ij}^E} \right) = \delta_j^E + \mathbf{Z}_i' \beta_Z^E, E = P, T \tag{8}$$

where $\mathbf{Z}_i$ are characteristics of student i; and δ_j is the effect of school j on the student's probability of failing.

Student-specific variables include each student's age and gender, his location relative to the city center as a proxy for whether the student lives in an urban or rural area, as well as attributes of the student's parish, such as educational attainment and income. We also include interactions between some of these variables and school characteristics such as instructor experience. For the on-road exam fail probability regression, we also include the examiner's experience and interactions of student attributes with examiner experience.

6.3 Likelihood function

We estimate the joint school choice, exam outcome model via maximum likelihood. Assuming that ε_{ij} in (5) is an i.i.d. type-1 extreme value random variable, the probability that student i chooses to attend school j, $c_{ij} = 1$, is conditional on the student's type h (inattentive or sophisticated), and can be written as

$$\Pr(c_{ij}=1|h) = \frac{\exp\left\{\mathbf{X}_j'\beta_X + \mathbf{Z}_i'\beta_Z + f\left(D_{ij}\right) - \alpha\left(p_j^u + \mathbb{I}_{\{h=s\}}\sum_{E=P,T}\hat{\pi}_{ij}^E p_j^{a,E}\right) + \xi_j\right\}}{\sum_{k=1}^{J}\exp\left\{\mathbf{X}_k'\beta_X + \mathbf{Z}_i'\beta_Z + f\left(D_{ik}\right) - \alpha\left(p_j^u + \mathbb{I}_{\{h=s\}}\sum_{E=P,T}\hat{\pi}_{ij}^E p_j^{a,E}\right) + \xi_k\right\}} \tag{9}$$

Conditional on type h, we assume that the school choice proabilities and the fail propensities are independent; as a result, the likelihood of student i failing exam E is given in Equation 8. The joint likelihood of observing the student's school choice and his two exam outcomes thus equals the product of the two likelihoods, for a student of type h:

$$\Pr(\lambda_{ij}^P = \lambda^P, \lambda_{ij}^T = \lambda^T, c_{ij} = 1|h) = P_{ij|h} = \prod_{E=P,T}\hat{\pi}_{ij}^{E\,\lambda^E}\left(1-\hat{\pi}_{ij}^E\right)^{(1-\lambda^E)} \times \frac{\exp(\hat{u}_{ij} - \alpha\mathbb{E}p_{ij|h})}{\sum_{k=1}^{K}\exp(\hat{u}_{ik} - \alpha\mathbb{E}p_{ik|h})} \tag{10}$$

Since we do not observe any student's degree of sophistication, given by h, we estimate a finite mixture model in which the probability of belonging to each of two possible segments is given by π_s and $\pi_m = 1 - \pi_s$, with π_s a parameter to be estimated. Thus, the unconditional likelihood of observing the full sample's school choices and exam outcomes equals

$$L = \prod_{i=1}^{N}\left(\pi_m P_{ij|\{h=m\}} + \left(1-\pi_m\right)P_{ij|\{h=s\}}\right). \tag{11}$$

6.4 Parameter Identification

So far, we have established that there are consumers who make their school choice with limited account of the add-on's price. The survey evidence in Section 5 suggests that this could be for several reasons, including underestimating either the price of the add-on or the probability of failing. The objective of the empirical model is not to identify the underlying causes of inattention present in the data (which is difficult with observational data only), but rather to estimate the proportion of inattentive consumers. Here, we discuss how the empirical choice model is able to identify two segments of consumers with distinct preferences.

First, consider only inattentive consumers (i.e. $\lambda = 0$) who base their school choice purely on each school's upfront price. The cost of a course of driving instruction then varies only across firms, but not by student at a given school. With no variation in prices across students, we rely on the variation in prices across schools and markets illustrated in Table 2 in order to estimate the students' price responsiveness.

Now consider the case with heterogeneity in consumers' decision processes such that inattentive students choose schools based on base prices alone, and sophisticated students select based on the total price. If student types were observed, we could account for such differential choice processes by varying the utility across students, with some students ignoring the add-on price. Since this is not the case in our data, we use a modeling approach akin to the commonly used mixed logit models with a discrete mixture of types (also known as latent class choice models). There are a few differences in our approach that are worth noting.

In common latent class approaches, a probabilistic model is used to allocate consumers to segments that represent different unobserved tastes in the sample. In our case, we furthermore allow students' decision criteria to differ across segments. That is, the probabilistic "allocation model" assigns consumers to each class based on whether they are more likely to choose between schools by taking into account the upfront price only or both the upfront and add-on prices. The use of this setup is intuitively appealing given that our economic model has two types (or classes) of consumers making decisions according to different decision processes.

What type of variation is then needed in the data in order to achieve identification of each segment's size? Intuitively, this problem reduces to understanding what kind of variation would allow estimation of the share of consumers who make choices that are more consistent with the upfront price, instead of the expected add-on price.

A basic requirement for differentiating between upfront and total expected prices as drivers of school choice would thus be observing variation in expected add-on prices within each student's municipality that is not perfectly correlated with upfront prices. Note that variation in expected add-on prices reflects both variation in add-on prices per se, which we demonstrate in Table 2, and variation in students' fail probabilities across schools. Identification of the share coefficient thus stems in part from choices made by individuals who have sufficiently high probabilities of failing such that they prefer to bypass schools with low upfront prices (which are preferred by inattentive students), in favor of schools that have lower expected total prices, all else equal. The results in Table 10 suggest that fail probabilities vary with student attributes such as gender and age; we further reject that the included school fixed effects are jointly insignificant in explaining exam outcomes in either of the specifications. For the chosen schools, the coefficient of variation in predicted fail probabilities based on the estimated models in Table 10 is 0.43 and 0.33 for the incidence of failing the theory and on-road exams, respectively, and rises to 0.54 and 0.49, respectively, once we include predicted fail probabilities for those schools in the student's municipality that the student ultimately did not choose. These two sources of variation introduce variation in expected add-on prices across students and schools.

To investigate the correlation between upfront and add-on prices, we compare the ranking of all schools in a student's municipality based on upfront prices and based on total prices. For over 35% of the students in the data, there is no significant statistical association (measured by the Spearman correlation coefficient) between the ranks of a school's upfront and total prices.

The estimated segment-share coefficient then represents the extent to which such variation between upfront and total prices translates into variation in choices. Descriptively, such variation is present in the data; 33.4% of students choose the school that is the lowest priced in its market based on either the upfront or the total price; of these 36.2% choose a school that is lowest in upfront, but not total, price; 13.51% choose a school that is lowest in total, but not upfront, price; and 50.29% choose a school that has both the lowest upfront and total price. Apart from price, the most important determinant of school choice is distance to the school; 44.1% of students choose the closest school, which is also the lowest priced along either of the price dimensions for 31.8% of those students. We investigate the separate explanatory power of upfront and total prices further in single-type school choice models below.

6.5 Results and Discussion

Preliminary model results for a subset of markets are depicted in Tables 10 and 11. These result from estimating the fail probabilities separately, and holding the estimated parameters fixed in estimating the school choice model. A shortcoming of these results is thus that the fail probabilities do not vary by student type. We begin with a discussion of the fail probability models depicted in Table 10. Fail propensities for both exams increase in the student's age, an effect that is more pronounced for female students. On average, however, female students are less likely to fail the theory exam, while they are more likely than male students to fail the on-road exam. The probability of failing the theory exam, but not the on-road exam declines significantly in educational attainment. Possibly correlated with educational attainment at the student-level, we also find that students living at a larger distance from the city center, and thus in more remote areas, have higher fail rates, and in particular if they are older, which again is not true for the on-road exam. Male students, more so than female students, benefit from instructor experience in lowering their theory pass rates. The examiner's experience contributes in lowering both the overall, and even more so, the female propensity of failing the on-road exam.

=========================

Insert Table 10 about here

=========================

We also use the exam outcome models to investigate descriptively whether there is evidence of a price discrimination motive for firms in setting repeat course fees. For the relative high prices of repeat courses compared to base course fees to reflect price discrimination, we would expect to see a correlation between the student's price sensitivity and their add-on demand, that is their likelihood of failing the theory or on-road exam. If more price sensitive students are more likely to pass the exam and thus have low add-on demand, firms would optimally respond with high repeat course prices. We include two measures of price sensitivity: first, we construct a measure of revealed price sensitivity that classifies students who choose a school that is not their closest, but charges the cheapest base course fee in the municipality. Second, we consider whether students who live further from the school, and thus have a higher travel cost, might engage in higher effort to pass the exam to reduce the full cost of attending school. To avoid the element of school choice to contaminate the distance measure, we focus on students who reside in monopoly markets and thus did not have a school choice to make. This is not possible for the first price sensitivity proxy, which relies directly on comparing the student's school choice to other alternatives. In Table 10, specifications (2), (3), (5) and (6) contain these additional variables. Neither revealed price sensitivity nor distance to the school correlate significantly with the propensity of failing either exam.

Table 11 contains estimates for several specifications of the school choice model. In specifications (1) and (2) we investigate how different components of the student's price affect his school choice. We decompose each student's expected price into his upfront price and expected repeat fees and enter these two prices separately into the utility function. This separation addresses the possibility that the observed price patterns in the data purely reflect differences in the students' price sensitivities across the two services, but not differences in attention across student types.

=========================

Insert Table 11 about here

=========================

Since we include the upfront price, a school-level covariate, we cannot include school fixed effects to control for unobserved heterogeneity across schools that might be correlated with prices. Instead, we include pertinent school attributes. The specification school attributes only, including total expected prices instead of the two components, is similar to the one with school fixed effects (specifications (1) and (3)), suggesting that much of the heterogeneity is observed.

The results for the first two specifications allow two conclusions: First, while slightly larger in absolute value, the effect of add-on prices on utility is not statistically significantly different from the effect of the upfront price; thus, price responsiveness is not more pronounced for repeat fees than it is for base prices. Second, the estimated coefficients remain largely unchanged, both for price and for the remaining school choice determinants, when we combine the two price components into a single expected price (specification (1)).

A downside to the single-type specifications is that we do not allow for unobserved heterogeneity in the estimated coefficients. We, therefore, now turn to specifications (4) through (6), which estimate discrete mixture models with two consumer types. The displayed specifications allow for heterogeneity in the price coefficient only. We also estimated alternative specifications that, in addition, allowed for heterogneity in both the distance covariates and the school fixed effects. The random coefficients on prices are robust to the addition of heterogeneity in other variables – that is, the result that inattentive students are the more price sensitive obtains when we allow for heterogeneity in other covariates. Moreover, inattentive students are less sensitive to distance, choosing a school that is farther away more frequently and being less likely to choose the closest school to their homes.

Similar to specifications (1) through (3), the two-type models differ in the way price enters into the utility function. Specifications (4) and (5) assume that utility depends only on expected prices for both market segments, but differ in that specification (4) includes school controls while specification (5) includes school fixed effects. These specifications thus assume that all consumers in the market are sophisticated, but differ in their price sensitivities, which we find to be the case for the two-type specifications.

Specification (6) then replaces the expected price with the base price, as in the model of inattentive consumer behavior in Section 2. Interestingly, we estimate a larger price coefficient in absolute value for the inattentive segment, suggesting that inattentive students are more price sensitive despite considering only the base price – rather than the higher expected price – at the time of their initial school choice. The estimated share parameter implies that approximately one quarter of students fall into the first (inattentive) market segment. The size of the segment we estimate is similar in size to the smaller segment in specifications (4) and (5) (there we estimate the smaller segment's size to be slightly higher at approximately 30 percent), which is similarly the more price responsive. Lastly, note that model (6) has the lowest likelihood value across specifications. The results thus suggest that observed consumer responses to the firm's chosen upfront and add-on prices is consistent with a sizable number of students who do not base their school choice on a comparison of their full expected fees at the different schools, but only on a comparison of the initial upfront payment.

7 Conclusion

In this paper, we have modeled a market in which firms take advantage of locked-in inattentive consumers by charging high fees in add-on markets. When consumers are inattentive to their demand for add-on products and service and face high switching costs, firms have the dual incentive to set 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. The model, hence, predicts upfront (add-on) prices that (do not) vary with the market structure a firm faces. The model also implies that all base prices depend on the proportion of consumer types in the market and the type-specific probability of requiring the 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 repeat fees, but also charge significantly higher markups in the add-on than in the upfront market. The evidence corroborates the model predictions that the upfront prices vary with the level of competition whereas add-on prices do not.

Our research benefits from a highly detailed data snapshot of the driving school industry, which gives us the ability to address additional questions of interest, most notably who the typical inattentive consumer is and which consumer type participates most in the add-on market, implicitly cross-subsidizing the other. We specify a two-type mixture model of demand that empirically pins down the proportion of student types that is consistent with the observed school choices under schools' chosen base and add-on prices. That about one-quarter of students is inattentive points to schools' strategic exploitation of this subset of students who do not anticipate their need for the add-on at their initial purchase occasion. As a next step, we plan to estimate the student fail probabilities jointly with the choice probabilities as well as to estimate a version of the model using concomitant variables, whereby we correlate the implied type share parameter with student attributes.

Evidence that speaks to the question of who subsidizes whom in a market such as ours is of significant normative policy interest to regulators of firm pricing behavior. In the case of policies under consideration by the IMTT, possible regulations range from requiring schools to inform students about typical propensities of failing the different exams to releasing price information to directly or indirectly regulating prices in the add-on market. As a next step, hence, we plan to analyze the attributes of schools whose market shares are specifically increased by the presence of inattentive students. By using the estimates of our empirical model to compare the case in which only some student consider the add-on to the case in which all students consider the add-on, we will be able to draw conclusions about those schools that benefit from having inattentive students in the market. Another avenue for future research is to implement an empirical pricing model in our setting, allowing us to compare counterfactual price predictions under varying assumptions about the student types in the 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.

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.

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

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.

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

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.

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

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.

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.

Shulman, J. D. and Geng, X. (2012). Add-on pricing by asymmetric firms. Forthcoming in Management Science.

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.

Figures and Tables

Figure 1: Driving Schools by Municipality, Mainland Portugal

(a) All municipalities

(b) Available sample of municipalities

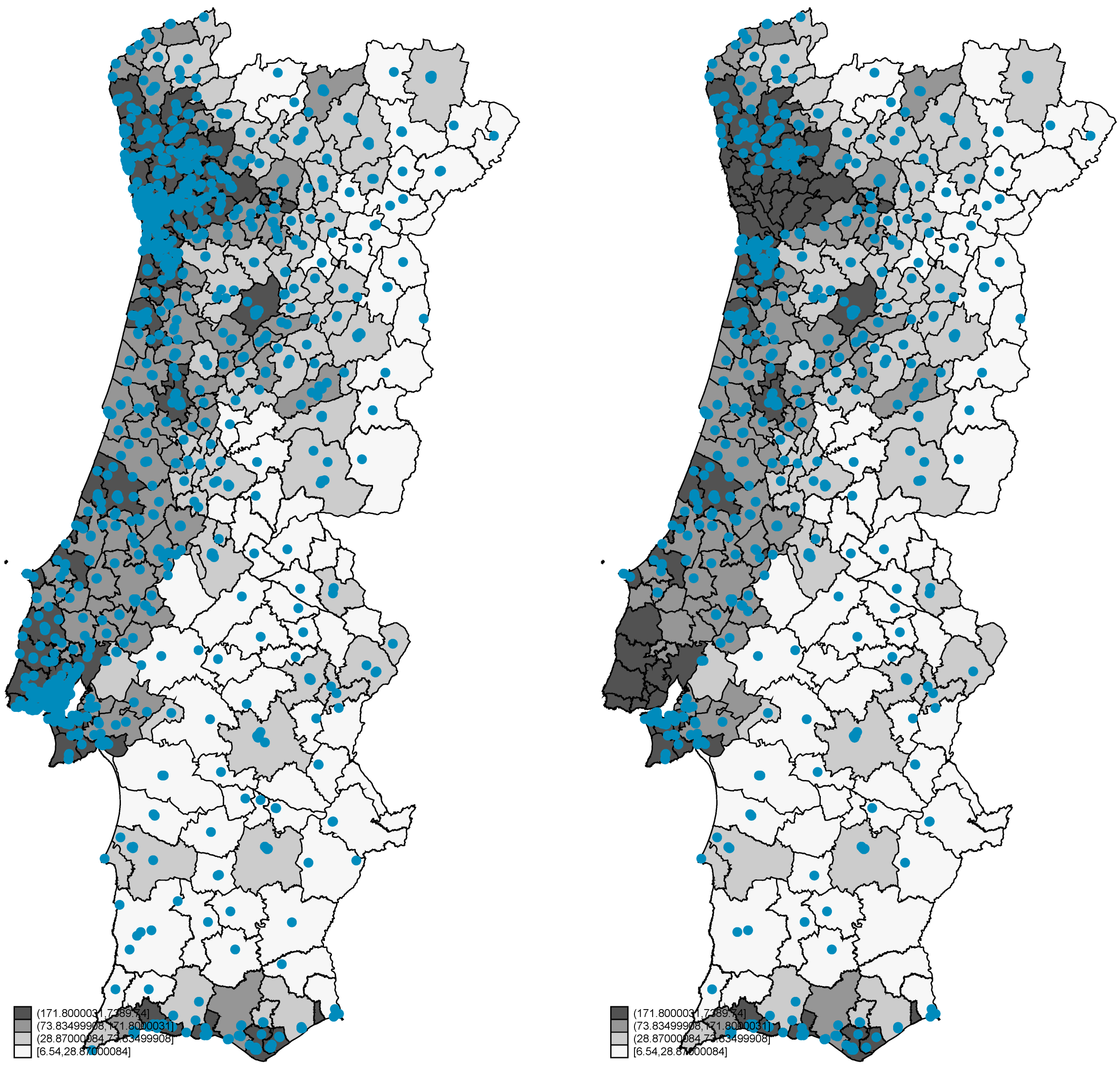


Figure 2: Number of On-road and Theory Exams Taken, by Student

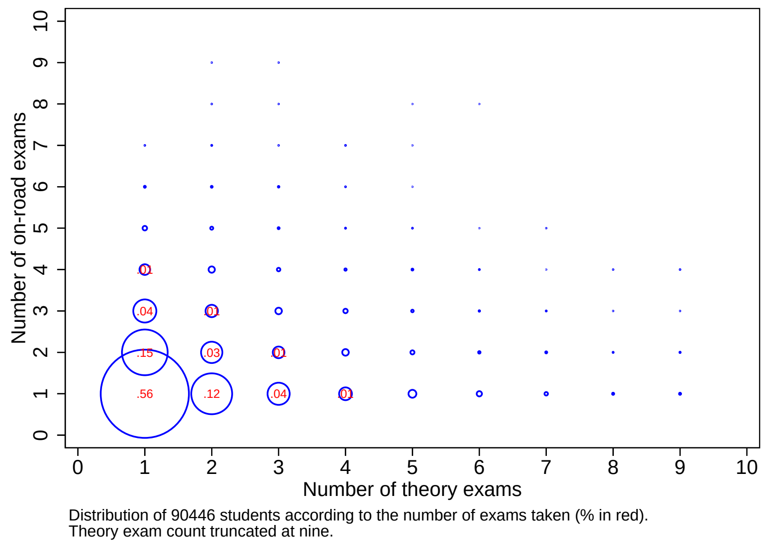


Figure 3: Distribution of Lerner Indices in Base and Repeat Course Markets, by School (%)

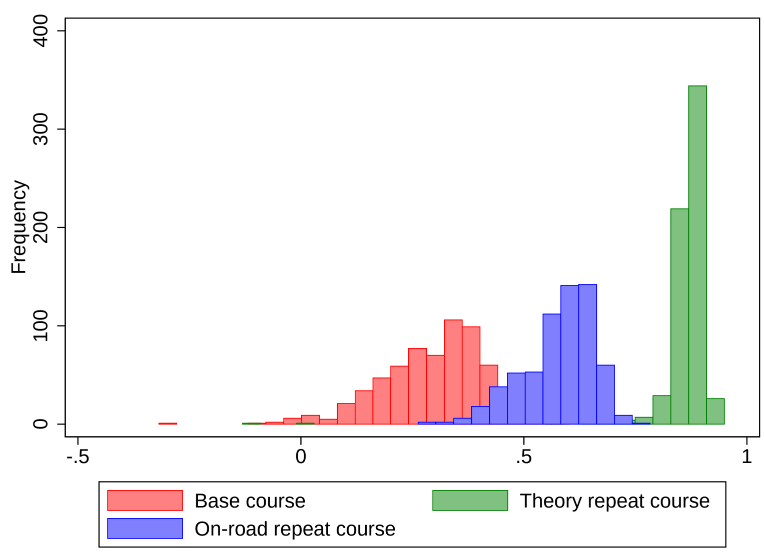
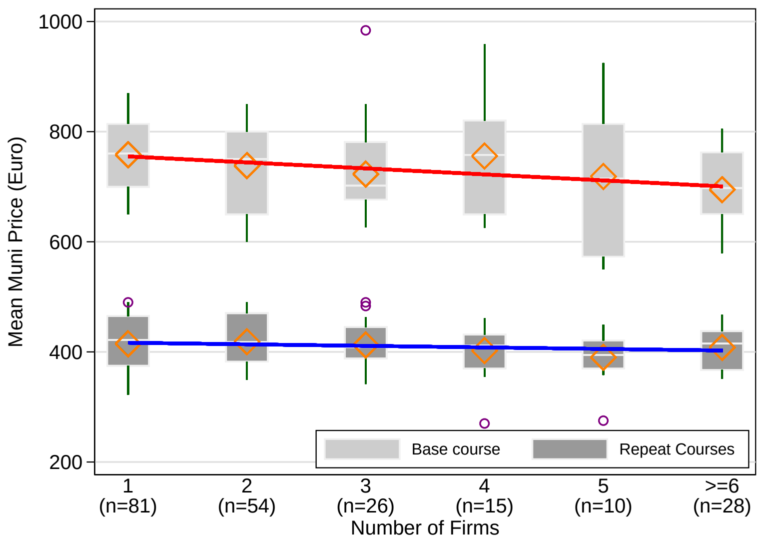


Figure 4: Average Municipality Prices by Number of Schools



Note: Interquartile range of municipality prices conditional on number of schools. Diamonds indicate mean price level.

Figure 5: Student Questionnaire

QUESTIONÁRIO QUESTIONNAIRE

No âmbito de um estudo sobre ensino e exames de condução agradecemos que preencha o seguinte questionário. We are conducting a study concerning drivers' education and testing and would appreciate if you could answer the following questions

Preencha o que for necessário ou assinale com um (X) a opção adequada
Please fill in the blank spaces or mark your selection with an X

1. Foi a primeira vez que fez exame teórico de condução? ☐Sim Yes ☐Não No
Was this the first time you took the theory exam?

2. O exame foi mais difícil do que esperava? ☐Sim ☐Não
Was the exam harder than you were expecting it to be?

3. Acha que vai passar no exame teórico que acabou de realizar? ☐Sim ☐Não
Do you think you will pass the exam you just took?

4. Caso reprove, vai ter de pagar à sua escola de condução para repetir o exame teórico?
In case you fail, will you have to pay your school to retake the theory exam?
☐Não Sei Don't know ☐Sim Yes ☐Não No

A) Aproximadamente quanto? (se não souber deixe em branco) ________ Euros
Approximately how much? (If you don't know please leave it blank)

Sei quanto é porque: ☐ Perguntei na escola ☐ A escola informou-me ☐ Outro: ________
I know how much it is because: I asked the school The school informed me Other

B) Quanto acha que você *devia* ter de pagar para poder repetir o exame? ________ Euros
How much do you think you would have to pay to be retake the exam?

5. Assinale com um (X) se cada situação descrita é VERDADEIRA ou FALSA
Mark with an (X) whether each of the following situations is True or False.

	VERDADE True	**FALSO** False
Fiz mais de 50 testes no computador como preparação para hoje As a preparation for today I did more than 50 computer (at home) tests	☐	☐
Fui fazendo testes à medida que ía assistindo às aulas teóricas I did practice tests during the entire time I was taking theory lessons	☐	☐
Usei o livro de código para perceber melhor os meus erros nos testes I used the "Rules of the Road" book to better understand my mistakes on those tests	☐	☐
Fui a mais aulas teóricas do que o mínimo exigido I attended more theory lessons than the minimum required	☐	☐
Preparei a matéria das aulas teóricas *antes* de ir assistir às aulas I prepared the lessons' material before attending the lessons	☐	☐
Tirei apontamentos nas aulas teóricas I took notes during the theory lessons	☐	☐
Tirei dúvidas com o instrutor várias vezes I asked the instructor for clarification several times	☐	☐
Acho que não é preciso estudar muito para o exame teórico I don't think one needs to study very hard for the theory exam	☐	☐

6. Quantas aulas práticas (de condução) já completou até à data de hoje?
How many practice lessons (on-the-road) have you completed before today? ________ Aulas Lessons

7. Acha que vai passar à primeira no exame prático (de condução)? ☐Sim ☐Não
Do you think you will pass the driving exam at first try?

8. Data de nascimento: ____/____/____
Date of birth

9. Sexo: ☐Feminino ☐Masculino
Gender

10. (OPCIONAL) Caso não se importe de ser contactado posteriormente para perguntas adicionais referentes ao exame por favor indique o seu endereço de email
(OPTIONAL) Please provide your email address if we could contact you with additional questions regarding the exam.
E-mail: ____________________

Obrigado pela sua colaboração!
Thank you very much!

Table 1: Summary Statistics, Municipality Characteristics, Sample Markets ($N = 214$)

	Mean	Std Dev	Q25	Med	Q75
Number of schools	3.224	3.271	1.000	2.000	4.000
Number of students	422.645	560.059	98.000	236.500	487.000
Population (000)	27.158	32.206	7.874	15.700	29.749
Share of population in urban parishes	80.014	21.989	66.180	87.796	100.000
Mean per-capita income (€)	861.254	113.706	772.700	867.400	919.900
Share with higher education (parish)	12.789	5.173	8.925	11.300	15.308
Number of parishes in muni	24.814	20.969	8.000	18.000	31.000
Number of car repair shops	62.030	55.415	19.000	40.000	90.000
Average monthly rent (€)	137.099	31.673	116.100	140.200	159.444
Share non-residential buildings	8.858	3.094	6.780	8.680	10.400

Note: Relevant statistics are student weighted. Mean income is monthly wage income of full-time employees.

Table 2: Summary Statistics, School and Examiner Characteristics

	Mean	Std Dev	Q25	Med	Q75
Number of students	142.211	83.677	87.000	122.000	176.000
Number of instructors	5.751	3.759	3.000	5.000	7.000
Instructor experience (yr)	8.044	5.417	4.202	6.458	10.214
Number of vehicles	3.822	2.359	3.000	3.000	4.000
Median weight of fleet cars (000 kg)	1.195	0.105	1.130	1.185	1.261
Distance to exam center (km)	21.142	17.893	5.619	17.883	30.081
Number of examiners per center	8.054	4.941	4.000	7.000	12.000
Examiner experience (exams/month)	57.892	18.914	48.308	57.131	69.944
Distance to IMTT office (km)	23.369	17.481	10.738	20.620	35.623
Price, base course	705.900	110.424	625.000	700.000	775.000
Price, theory course repeat	129.080	26.288	110.330	130.000	150.000
Price, on-road course repeat	275.209	47.076	250.000	275.880	305.000

Note: Observations: 636 schools, 26 exam centers, and 185 examiners. All statistics are student weighted.

Table 3: Summary Statistics, Student Attributes ($N = 90,446$)

	Mean	Std Dev	Q25	Med	Q75
Age at time of theory exam (yr)	21.876	6.771	18.349	19.072	21.812
Gender (1=F,0=M)	0.510	0.500	0.000	1.000	1.000
Distance to city center (km)	4.644	3.872	1.482	3.780	6.824
Distance to school (km)	4.874	6.007	1.007	2.875	6.268
Theory exams taken (no)	1.371	0.812	1.000	1.000	1.000
On-road exams taken (no)	1.360	0.688	1.000	1.000	2.000
Pass rate, first theory exam (%)	0.760	0.427	1.000	1.000	1.000
Pass rate, first on-road exam (%)	0.732	0.443	0.000	1.000	1.000
Time to completion (days)	250.116	150.527	143.000	211.000	315.000
Choice is closest school (1=Y,0=N)	0.464	0.499	0.000	0.000	1.000

Table 4: Relative Prices and Estimated Markups for Base and Repeat Courses

	Mean	Std Dev	Q25	Med	Q75
Ratio, repeat to base fees	0.585	0.117	0.500	0.587	0.667
Markup, base course	203.590	111.777	124.302	199.200	279.150
Markup, theory course repeat	112.712	26.111	95.000	113.797	133.000
Markup, on-road course repeat	161.369	45.908	132.427	162.689	194.120
Effective markup	313.802	121.180	220.454	309.519	400.541
Percent markup, base course (%)	0.272	0.124	0.196	0.286	0.361
Percent markup, theory course repeat (%)	0.864	0.069	0.850	0.875	0.889
Percent markup, on-road course repeat (%)	0.575	0.080	0.533	0.589	0.630
Percent effective markup (%)	0.338	0.097	0.276	0.349	0.414

Note: The effective markup is the average total markup across a school's students.

Table 5: Regression Models of Base Course Prices

	OLS		IV	
	(1)	(2)	(3)	(4)
Number of firms (N)	−9.524**		−10.874**	
	(4.190)		(5.293)	
($N = 2$) Y/N		−28.155*		−26.390
		(15.818)		(17.938)
($N = 3$) Y/N		−31.531		−32.467
		(20.077)		(23.412)
($N = 4$) Y/N		−55.937**		−56.327*
		(24.269)		(31.832)
($N = 5$) Y/N		−69.811**		−67.881***
		(30.334)		(25.442)
($N \geq 6$) Y/N		−85.101***		−84.558**
		(29.120)		(33.311)
% fail both exams	−2.461**	−2.581**	−2.474**	−1.683*
	(1.105)	(1.115)	(1.086)	(0.863)
Number of instructors	0.786	1.669	0.686	2.125
	(1.292)	(1.235)	(1.239)	(1.372)
Median weight of fleet cars (kg)	−0.013	−0.002	−0.013	0.020
	(0.035)	(0.036)	(0.034)	(0.028)
School ≤ 100m to high school	48.488	50.206	48.319	41.008
	(39.767)	(39.137)	(39.236)	(31.304)
School distance to city center	3.827***	4.445***	3.862***	3.256***
	(1.240)	(1.315)	(1.214)	(0.798)
% of population in urban parishes	−0.650*	−0.604*	−0.636*	−0.407*
	(0.346)	(0.318)	(0.348)	(0.360)
Per-capita income	0.109	0.152**	0.107	−0.084*
	(0.076)	(0.075)	(0.076)	(0.046)
Number of car repair shops	−0.552**	−0.588**	−0.528*	0.729
	(0.264)	(0.286)	(0.278)	(0.515)
Population (000)	0.850**	0.326	0.951**	−0.511
	(0.400)	(0.367)	(0.417)	(0.741)
Adjusted R^2	0.360	0.369	0.360	
1^{st} stage partial R^2			0.689	
1^{st} stage F statistic			16.614	
Observations	636	636	636	636

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Municipality-level clustered standard errors in parentheses.

Note: Specifications include region fixed effects and additional, statistically insignificant, controls (instructor experience and its square, distance to exam center, and a coastal area indicator). Models (3) and (4) use pre-deregulation firm count, distance to the closest IMTT office, average rent and commercial building share as instruments.

Table 6: Regression Models of Repeat Course Price

	OLS		IV	
	(1)	(2)	(3)	(4)
Number of firms (N)	−1.063		−3.972	
	(2.249)		(3.216)	
($N = 2$) Y/N		3.157		7.672
		(10.706)		(10.690)
($N = 3$) Y/N		−5.246		5.045
		(12.440)		(14.415)
($N = 4$) Y/N		−4.438		−2.108
		(18.767)		(17.097)
($N = 5$) Y/N		−18.969		−19.788
		(15.643)		(17.134)
($N \geq 6$) Y/N		8.083		14.587
		(15.676)		(24.098)
% fail both exams	−0.444	−0.468	−0.473	0.092
	(0.583)	(0.574)	(0.585)	(0.716)
Number of instructors	4.210***	4.330***	3.993***	3.164***
	(0.839)	(0.863)	(0.818)	(0.935)
Median weight of fleet cars (kg)	0.078***	0.076***	0.078***	0.049***
	(0.023)	(0.023)	(0.023)	(0.023)
School ≤ 100m to high school	55.726***	55.475***	55.362***	50.728**
	(20.359)	(20.991)	(19.926)	(25.737)
School distance to city center	1.776**	1.993**	1.852**	1.142*
	(0.848)	(0.897)	(0.832)	(0.604)
% of population in urban parishes	−0.572***	−0.521***	−0.541***	−0.634***
	(0.169)	(0.171)	(0.167)	(0.217)
Per-capita income	0.012	0.011	0.006	−0.017
	(0.040)	(0.041)	(0.040)	(0.038)
Number of car repair shops	0.487***	0.453***	0.540***	0.298
	(0.122)	(0.136)	(0.126)	(0.350)
Population (000)	−0.459*	−0.621***	−0.242	−0.431
	(0.247)	(0.155)	(0.328)	(0.351)
Adjusted R^2	0.145	0.153	0.141	
1st stage partial R^2			0.689	
1st stage F statistic			16.614	
Observations	636	636	636	636

* p < 0.1, ** p < 0.05, *** p < 0.01. Municipality-level clustered standard errors in parentheses.
Note: See notes to Table 5.

Table 7: Expectations of Financial Repercussions of Failing Theory Exam, Survey Respondents

	Share of Respondents		
	All	Female	Male
No	4.091	3.070	5.189
Do not know	17.500	14.912	20.283
Yes, but do not know how much	26.364	25.439	27.358
Yes, and know how much	52.045	56.579	47.170**

Note: Breakdown of responses of 440 survey participants to question "In case you fail, will you have to pay your school to retake the theory exam?". P-values for test of equality of male and female shares of respondents indicated as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Comparison of Expected and Actual Exam Outcomes, Survey Respondents

	Theory		On-Road	
	Pass Rate	Mean Difference, Expectation less Incidence of Passing	Pass Rate	Mean Difference, Expectation less Incidence of Passing
Overall	0.725	−0.034	0.766	0.100***
Female	0.711	−0.083**	0.739	0.075*
Male	0.741	0.019	0.792	0.126***
Informed	0.733	−0.035	0.786	0.089***
Uninformed	0.695	−0.032	0.683	0.143**

Note: Informed students: Expect to pay for exam retake = Y. P-values indicated as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The share of female and informed students for the theory (on-road) exam comparison is 51.82% and 78.41% (50.31% and 80.31%). There is no statistically significant difference at the 5% level between female and male or informed and uninformed students' pass rates and between informed and uninformed students' outcome expectations. Male students have higher pass rate expectations than female students for both the theory and the on-road exams at the 1% level.

Table 9: Exam Preparation by Student Type, Survey Respondents

	Type 1	Type 2	p-Value
Active Preparation (Index)			
Type 1 = Informed	−0.034	0.060	0.454
Type 1 = Female	−0.021	−0.005	0.877
Passive Preparation (Index)			
Type 1 = Informed	0.041	−0.208	0.046
Type 1 = Female	0.216	−0.253	0.000
Overall Preparation (Share)			
Type 1 = Informed	0.654	0.637	0.502
Type 1 = Female	0.680	0.620	0.005
Study Little Y/N[1]			
Type 1 = Informed	0.120	0.140	0.625
Type 1 = Female	0.073	0.178	0.001

Note: Based on responses to seven questions on extent of exam preparation, e.g., "read book YN" [passive] or "interact with instructor YN" [active]. "Overall preparation" refers to the share of these preparation categories the student engaged in. "Study Little Y/N" is an indicator response to the question of "Do you feel you did not need to study much in preparation for the exam?".

Table 10: Conditional Logit Model of Propensity of Failing Exams

	Theory Exam			On-Road Exam		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−0.475***	−0.462***	−0.409*	0.519***	0.464***	0.847***
	(0.067)	(0.071)	(0.222)	(0.083)	(0.087)	(0.297)
Age	0.015***	0.016***	0.022***	0.024***	0.024***	0.033***
	(0.003)	(0.003)	(0.007)	(0.003)	(0.003)	(0.008)
Female*Age	0.015***	0.015***	0.010	0.023***	0.024***	0.011
	(0.002)	(0.002)	(0.009)	(0.002)	(0.003)	(0.009)
Share with higher education (parish)	−0.644***	−0.625***	−2.447**	0.187	0.217	−0.257
	(0.170)	(0.173)	(1.215)	(0.157)	(0.160)	(1.240)
Distance to city center	0.006*	0.005	0.002	0.005	0.005	0.034*
	(0.003)	(0.003)	(0.018)	(0.003)	(0.003)	(0.019)
Distance to city center*Age≥21	0.006*	0.006	0.008	−0.006	−0.005	−0.032**
	(0.004)	(0.004)	(0.015)	(0.004)	(0.004)	(0.016)
Instructor experience*Female	0.006*	0.006*	0.008	−0.003	−0.004	−0.006
	(0.003)	(0.003)	(0.010)	(0.003)	(0.003)	(0.010)
Examiner experience				−0.002***	−0.003***	0.007**
				(0.001)	(0.001)	(0.003)
Examiner experience*Female				−0.003***	−0.002**	−0.002
				(0.001)	(0.001)	(0.004)
Price Sensitive Y/N		0.014			0.003	
		(0.034)			(0.034)	
Distance to school			−0.002			−0.022
			(0.017)			(0.017)
Observations	90446	84618	5828	90446	84618	5828
Log-likelihood	−46843.675	−43677.546	−3014.810	−46736.242	−43608.019	−2986.512

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: Price Sensitive Y/N indicates students whose chosen school is the lowest priced in the market, but not the closest to their home. School fixed effects included in all models.

Table 11: School Choice Model: Preliminary Results

	1-Type			2-Types		
Variables	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Share (π_1)	——	——	——	−0.7857*	−0.9531**	−1.0677***
				(0.4160)	(0.3869)	(0.3949)
Upfront price (α_1)		−30.0456***				−82.3891***
		(3.1796)				(12.1207)
Expected add-on price		−36.3258***	−37.3828***			
		(3.8932)	(4.6649)			
Expected price (α_1)	−30.3315***			−57.6572***	−70.8361***	
	(3.1514)			(8.0111)	(9.3543)	
Expected price (α_2)				−26.7551***	−39.4712***	−47.8560***
				(4.8583)	(5.8862)	(10.1984)
Distance	−10.0190***	−10.0513***	−9.8275***	−10.8074***	−10.6086***	−10.4420***
	(0.7741)	(0.7746)	(0.7775)	(0.6885)	(0.6881)	(0.5424)
Distance^2	14.1295***	14.2360***	13.8343***	16.5583***	16.1440***	15.8499***
	(2.4352)	(2.4345)	(2.3251)	(2.3152)	(2.2489)	(2.2464)
Distance*female	1.1431**	1.1499**	1.1763**	1.0848**	1.1727**	1.2497**
	(0.5038)	(0.5019)	(0.5056)	(0.5293)	(0.5293)	(0.5047)
School Controls						
Instructor experience	6.7250***	7.2213***		8.5484***		
	(0.7726)	(0.7920)		(1.1006)		
Number of specialty courses	0.1269***	0.0796***		0.1502***		
	(0.0256)	(0.0290)		(0.0310)		
Number of vehicles	0.2939***	0.2774***		0.3743***		
	(0.0438)	(0.0431)		(0.0545)		
Age of school	−4.6264***	−4.9081***		−5.6160***		
	(0.4643)	(0.4786)		(0.6549)		
Distance to IMTT office	1.2965***	1.2395***		1.4586***		
	(0.2281)	(0.2282)		(0.2251)		
Distance to nearest university	−2.0957***	−2.1876***		−2.3872***		
	(0.4113)	(0.4212)		(0.4617)		
School fixed effects	N	N	Y	N	Y	Y
Log Likelihood	1,749.72	1,744.76	1,737.54	1,739.80	1,728.12	1,725.46
N	1,804	1,804	1,804	1,804	1,804	1,804

Note: The segment 1 and segment 2 share parameters equal $\theta_1 = \frac{\exp\{\pi_1\}}{1+\exp\{\pi_1\}}$ and $\theta_2 = 1 - \theta_2$, respectively. The implied segment 1 share values for specifications (iv), (v), and (vi) are 31.3, 27.8, and 25.6 percent, respectively.

Appendix

A Equilibrium in Illustrative Model in Section 2

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; inattentive 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 (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 sophisticates in the student population be $\pi \in (0,1)$ 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 period 2. There is a unique symmetric Nash equilibrium in which sophisticates engage in effort to lower their probability of failing to $\underline{\lambda}$ and, in period 1, schools charge an upfront price of

$$(p^u)^* = c^u + \frac{\sigma n}{n-1} - \left[(1-\pi)\overline{\lambda} + \pi\underline{\lambda}\right](\overline{p}^a - c^a)$$

while, in period 2, they charge an add-on price of

$$(p^a)^* = \overline{p}^a > c^a.$$

Consider first the school's choice of add-on price, p^a. In the second period, school j sells the add-on service to π sophisticated and $(1-\pi)$ inattentive students and earns profit of

$$\Pi_j = \pi\left[p^u_j - c^u + \underline{\lambda}\left(p^a_j - c^a\right)\right] D^s\left(p^u_{-j} - p^u_j + \underline{\lambda}(p^a_{-j} - p^a_j)\right) + \\ (1-\pi)\left[p^u_j - c^u + \overline{\lambda}\left(p^a_j - c^a\right)\right] D^m\left(p^u_{-j} - p^u_j\right), \tag{A.1}$$

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

Given that the add-on price, p^a, does not shift the demand of the inattentive 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 inattentive students through the decline in the upfront price and – provided $\overline{\lambda} > \underline{\lambda}$ – the revenue per inattentive student.[16]

The firm's upfront price p_u then maximizes:

$$\Pi_j = \left\{p^u_j - c^u + \left[\pi\underline{\lambda} + (1-\pi)\overline{\lambda}\right](\overline{p}_a - c^a)\right\} D^m\left(p^u_{-j} - p^u_j\right). \tag{A.2}$$

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 inattentive students' demand, D^m. Solving the first-order condition to the

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

firm's upfront pricing problem results in equilibrium prices of:

$$(p^u)^* = c^u + \frac{\sigma n}{n-1} - \left[(1-\pi)\overline{\lambda} + \pi\underline{\lambda}\right](\overline{p}^a - c^a)\,. \tag{A.3}$$

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 engaging in effort, $\left(\overline{\lambda} - \underline{\lambda}\right)\overline{p}^a$, exceed the cost of effort e. Otherwise, the optimal upfront price simplifies to:

$$(p^u)^* = c^u + \frac{\sigma n}{n-1} - \overline{\lambda}\,(\overline{p}^a - c^a)\,. \tag{A.4}$$

Equation (A.4) 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

$$(p^u)^* + \overline{\lambda}\overline{p}^a = c^u + \overline{\lambda}c^a + \frac{\sigma n}{n-1}. \tag{A.5}$$

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

$$\Pi_j = \left[p^u_j + \overline{\lambda}p^a_j - \left(c^u + \overline{\lambda}c^a\right)\right] D^s\left(p^u_{-j} - p^u_j + \overline{\lambda}(p^a_{-j} - p^a_j)\right) \tag{A.6}$$

In the absence of inattentive 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 the expected revenue in (A.5). Add-on pricing in the presence of inattentive students thus does not change the total amount of revenue 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.

Note also that with unavoidable add-ons and no cost to unshrouding prices, it can be shown that the firm is indifferent between shrouding and unshrouding. This result differs from the equilibrium derived in Gabaix and Laibson (2006) for the case of avoidable add-ons;[17] 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 how we compute marginal costs for the schools' three services.

The base course marginal cost comprises five cost components: the fees paid to the exam center for the theory and on-road exams (F^T and F^P, respectively), the cost of the instructional materials the school provides to the student (M), the instructor wages (W), and the cost of operating a fleet car (V) for a student's practice lessons. The theory repeat course generates as cost only the exam center fee for an examination. The on-road repeat course involves additional driving lessons with associated scaled-down wage and vehicle

[17]A number of authors, including Miao (2010), Kosfeld and Schüwer (2011), Heidhues et al. (2012), and Shulman and Geng (2012) have confirmed the results in Gabaix and Laibson (2006) in more general settings.

operating costs, in addition to the exam center's fee. Accordingly, we specify the marginal costs, MC^s, for service $s = \{U, T, P\}$ as:

$$MC^U = F^T + F^P + M + W + (700 + 2D)\,V \tag{B.1}$$

$$MC^T = F^T \tag{B.2}$$

$$MC^P = F^P + \frac{6}{33}W + \left(\frac{6}{33}\cdot 700 + 2D\right)V, \tag{B.3}$$

and compute each cost component as follows:

1. Exam adminstration fees (F^T and F^P) IMTT provided us with information on each of the 25 exam centers' fees for administering the theory and on-road exams. We use the administration fees of the exam center used by the school, or a weighted average of fees if the school uses multiple exam centers.

2. Instructional materials expenses (M) IMTT and ANIECA quote €10 for instructional materials (driver handbooks and CD-ROMs).

3. Instructor wages (W) Based on interviews with ANIECA and school representatives, an instructor's monthly salary ranges between €750 and €950. We assume that instructor salaries are proportional to mean monthly earnings in the school's municipality across municipalities. We gross up the salary to include 23.75 percent social security tax and a €3.4 per diem stipend.

   Since each student's base and on-road repeat courses include 32 and five driving lessons, respectively, we convert the monthly salary figure to an hourly basis based on ANIECA information on length of working days and number of days worked per month. We assume that schools incur only fixed, and not marginal, costs for the student's classroom time since schools rarely operate at capacity. The resulting marginal labor cost averages to €236.94 and only €43.08 for the base and repeat courses, respectively.

4. Vehicle operating costs $[(700 + 2D)\,V]$ We follow existing methodologies for computing a vehicle's user cost in Portugal, comprised of (i) fuel costs, (ii) depreciation costs, (iii) maintenance and repairs costs, and (iv) tire costs. For the base course, we scale this user cost per kilometer, V, by twice the return distance to the exam center plus the 700 kilometers that school owners state are covered during lessons; for the repeat course, the latter amounts to only 127 kilometers. The sources of data for the individual cost components are:

   - Fuel costs. We measure fuel costs as the average local price per liter of diesel fuel (obtained from the Direcção-Geral de Energia e Geologia on March 12, 2012 for each of the five lowest priced stations per municipality), times fuel consumption per kilometer. ANIECA and school owners state typical consumption rates of 6.36 liters per 100 km.
   - Depreciation costs. We follow existing methodologies whereby a vehicle depreciates fully by 8.4 years and has a purchase price of €25,000, on average. This, together with information on the average distance traveled by a fleet car per year, yields an estimate of cost-per-kilometer driven.
   - Maintenance and repair costs. We use public estimates of an average maintenance and repair cost of €4,000 over the car's service life, which we adjust to reflect fleet characteristics relative

Table B.1: Marginal Cost Components for All Services (€, $N = 636$)

	Mean	Std Dev	Q25	Med	Q75
Base Course					
Exam fees	53.022	9.141	45.000	49.000	59.000
Instructional materials	10.000	0.000	10.000	10.000	10.000
Instructor wages	235.629	18.277	220.274	237.242	246.648
Vehicle operating costs	194.896	3.504	192.310	195.519	197.602
Total	493.548	18.686	479.949	493.512	507.288
Theory Repeat Course					
Theory exam fee	16.368	1.868	15.000	15.384	17.000
On-road Repeat Course					
On-road exam fee	36.650	7.346	30.000	33.429	42.000
Instructor wages	42.842	3.323	40.050	43.135	44.845
Vehicle operating costs	32.923	2.950	30.339	33.410	35.346
Total	112.415	8.092	105.774	110.607	118.253

to the average car in the sample, and convert it to an estimate of cost per kilometer driven based on the annual distance traveled and a life of 8.4 years.

- Tire costs. We assume that the average car requires four new tires for every 40,000 kilometers at a cost of €70 per tire, which translates into an average tire cost of €.01 per kilometer

The vehicle operating cost requires an estimate of the annual distance covered per vehicle. Since the IMTT updates school fleet information only periodically, we employ the median across all fleet cars of 20,358 kilometers per vehicle-year for all schools. We use these inputs to calculate each school's marginal cost for its base and repeat courses as in Equation (B.1). Table B.1 summarizes the inputs into the cost calculation and the resulting totals.

As a robustness check, we recalculate marginal cost and markup using each school's data on kilometers covered per vehicle-year, rather than the median across cars. This alternative marginal cost measure averages to €502.30 for the upfront service and €113.84 for the on-road repeat course. The markup patterns in Figure 3 and regression results in Tables D.1 and D.2 are robust to using this alternative marginal cost.

C Nonlinear Specification Estimation Procedure

In this Appendix, we describe the full information maximum likelihood (FIML) procedure we use to control for the endogeneity of the discrete market structure indicators in the nonlinear price regressions describe in Section 4.4.

We assume that school i's price in market m is a function of observable market attributes, $\mathbf{X^1}$, market and school specific variables, $\mathbf{X^2}$, market structure indicators γ^P_{jm}, and a school and market specific unobservable,

ξ:

$$p^P_{im} = \alpha^P + \beta^{P,1}\mathbf{X}^1_m + \beta^{P,2}\mathbf{X}^2_{im} + \sum_{j=2}^{6} \gamma^P_{jm} \mathbb{1}_{\{\Pi^E_m = j\}} + \xi^P_{im} = f^P(\mathbf{X}^1_m, \mathbf{X}^2_{im}, \beta^{\mathbf{P}}) \tag{C.1}$$

We assume the unobservable can be decomposed into $\xi^P_{im} = \varepsilon^P_m + \eta^P_{im}$, where ε^P_m is a common error component shared by all schools in market m and η^P_{im} is the school-specific error term with $\eta^P_{im} \sim N\left(0, \sigma^2_{\eta,P}\right)$.

We specify the number of schools in market m, which ranges from 1 to 6 in the data, as an ordered probit model:[18]

$$\Pi^E_m = \begin{cases} 1 & \text{if } \alpha^E + \beta^{E,1}\mathbf{X}^1_m + \beta^{E,2}\mathbf{Z}_m + \varepsilon^E_m < \zeta^E_1 \\ j & \text{if } \zeta^E_j < \alpha^E + \beta^{E,1}\mathbf{X}^1_m + \beta^{E,2}\mathbf{Z}_m + \varepsilon^E_m < \zeta^E_{j+1} \text{ for } j = 2, .., 5 \\ 6 & \text{if } \alpha^E + \beta^{E,1}\mathbf{X}^1_m + \beta^{E,2}\mathbf{Z}_m + \varepsilon^E_m > \zeta^E_6, \end{cases} \tag{C.2}$$

where the parameter ζ^E_j implies a cutoff for the unobservable ε^E_m between $j-1$ and j schools and $\mathbf{Z}^E_m$ contains market-specific attributes that are excluded from the pricing equation.

We assume that ε^P_m and ε^E_m are distributed bivariate normal as follows:

$$\begin{pmatrix} \varepsilon^P_m \\ \varepsilon^E_m \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_P & \sigma_{EP} \\ \sigma_{EP} & 1 \end{pmatrix} \right) \tag{C.3}$$

The covariance terms allow for correlations in the market-level unobservables that give rise to endogeneity concerns, with σ_{EP} and σ_P parameters to be estimated.

In estimating the nonlinear system of equations, the contribution of the likelihood from market m equals:

$$L_m = \Pr\left(p^P_{im} = p_i \forall i, \Pi^E_m = j\right) = \Pr\left(\xi^P_{im} = p_i - f^P_i \forall i, \zeta^E_j - f^E < \varepsilon^E_m < \zeta^E_{j+1} - f^E\right) \tag{C.4}$$

where j is an index of the observed number of entrants, p_{im} is the observed price of school i, and $f^E(\mathbf{X}^1_m$, $\mathbf{Z}_{im}$, $\beta^{\mathbf{E}}$)= α^E+ $\beta^{E,1}\mathbf{X}^1_m$+ $\beta^{E,2}\mathbf{Z}_m$. This probability is given by the integral of the $2N+1$-dimensional normal distribution of ξ^P_{im} and ε^E_m with mean zero and variance-covariance matrix given by (where $\mathbf{I}$ is the identity matrix and $\mathbf{\Xi}$ is a matrix of ones)

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma^2_P \mathbf{\Xi}_{2N\times 2N} + \mathbf{I}_{2N\times 2N} & \sigma_{EP}\mathbf{I}_{2N\times 1} \\ \sigma_{EP}\mathbf{I}_{1\times 2N} & 1 \end{bmatrix} \tag{C.5}$$

over the surface defined by f^P and f^E and the cutoffs ζ^E_2 through ζ^E_6 that are consistent with observed prices and the observed number of entrants, respectively. Note that $\mathbf{\Sigma}$ results from stacking the $2N$ price equation errors $\xi^P_{im} = \varepsilon^P_m + \eta^P_{im}$ and the single market-level error, ε^E_m. We integrate out η^P_{im} to yield:

$$L_m = \int_{\zeta^E_j - f^E}^{\zeta^E_{j+1} - f^E} \int_{-\infty}^{\infty} \left[\prod_{i=1}^{N} \phi\left(p_i - f^P_i - \varepsilon^P_m\right) \right] \phi\left(\varepsilon^P_m, \varepsilon^E_m\right) d\varepsilon^P_m d\varepsilon^E_m, \tag{C.6}$$

where

$$g^P_i\left(\varepsilon^P_m\right) = \phi\left(p_i - f^P_i - \varepsilon^P_m\right)$$

is the standard normal pdf of η^P_{im} and $\phi\left(\varepsilon^P_m, \varepsilon^E_m\right)$ refers to the pdf of the bivariate normal distribution of $\left(\varepsilon^P_m, \varepsilon^E_m\right)$ given by C.3.

[18] We combine markets with six or more schools into the final category.

Conditioning on ε_m^P and integrating over ε_m^E results in:

$$L_m = \int_{-\infty}^{\infty} \left[\prod_{i=1}^{N} g_i^P \left(\varepsilon_m^P \right) \right] \times \left[\Phi_{\varepsilon_m^E | \varepsilon_m^P} \left(\zeta_{j+1}^E - f^E \right) - \Phi_{\varepsilon_m^E | \varepsilon_m^P} \left(\zeta_j^E - f^E \right) \right] \phi \left(\varepsilon_m^P \right) d\varepsilon_m^P, \tag{C.7}$$

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 and use a numerical optimizer to maximize the full likelihood:

$$L = \prod_{m=1}^{M} \left\{ \int_{-\infty}^{\infty} \left[\prod_{i=1}^{N} \frac{1}{\sigma_{\eta,P}} \phi \left(\frac{p_i - f_i^P - \varepsilon_m^P}{\sigma_{\eta,P}} \right) \right] \times \left[\Phi_{\varepsilon_m^E | \varepsilon_m^P} \left(\zeta_{j+1}^E - f^E \right) - \Phi_{\varepsilon_m^E | \varepsilon_m^P} \left(\zeta_j^E - f^E \right) \right] \phi \left(\varepsilon_m^P \right) d\varepsilon_m^P \right\}. \tag{C.8}$$

D Determinants of Markups

Here, we supplement the regression analysis in Section 4.4 with regression models of estimated school markups. These regressions allow us to verify that our primary results are not driven by cost differences between schools and across markets that are reflected in price. Table D.1 reports results for base course markups; Table D.2 for repeat course markups, with similar patterns to the effect of competition and fail rates as in Tables 5 and 6.

Table D.1: Regression Models of Base Course Markups

	OLS		IV	
	(1)	(2)	(3)	(4)
Number of firms (N)	−9.362**		−11.131**	
	(4.321)		(5.549)	
($N = 2$) Y/N		−34.762**		−33.206*
		(16.286)		(17.675)
($N = 3$) Y/N		−40.250*		−42.081*
		(20.472)		(23.881)
($N = 4$) Y/N		−63.935**		−65.831**
		(24.925)		(30.549)
($N = 5$) Y/N		−69.441**		−69.280***
		(30.606)		(24.910)
($N \geq 6$) Y/N		−88.446***		−88.391***
		(29.133)		(32.015)
% fail both exams	−2.506**	−2.623**	−2.524**	−1.648*
	(1.132)	(1.129)	(1.113)	(0.847)
Number of instructors	0.724	1.588	0.592	2.084
	(1.305)	(1.250)	(1.250)	(1.349)
Median weight of fleet cars (kg)	−0.027	−0.013	−0.027	0.011
	(0.035)	(0.036)	(0.035)	(0.029)
School ≤ 100m to high school	50.766	55.019	50.544	43.422
	(40.803)	(39.930)	(40.270)	(32.087)
School distance to city center	3.656***	4.225***	3.702***	3.101***
	(1.239)	(1.315)	(1.213)	(0.798)
% of population in urban parishes	−0.640*	−0.571*	−0.621*	−0.372
	(0.347)	(0.324)	(0.350)	(0.360)
Per-capita income	−0.011	0.031	−0.014	−0.016
	(0.077)	(0.074)	(0.077)	(0.048)
Number of car repair shops	−0.510*	−0.547*	−0.478	−0.614*
	(0.282)	(0.300)	(0.297)	(0.476)
Population (000)	0.789*	0.265	0.921**	0.307
	(0.414)	(0.378)	(0.431)	(0.719)
Adjusted R^2	0.326	0.336	0.326	
1st stage partial R^2			0.689	
1st stage F statistic			16.614	
Observations	636	636	636	636

* p < 0.1, ** p < 0.05, *** p < 0.01. Municipality-level clustered standard errors in parentheses.
Note: For a list of suppressed controls and instruments, see footnote to Table 5.

Table D.2: Regression Models of Repeat Course Markups

	OLS		IV	
	(1)	(2)	(3)	(4)
Number of firms (N)	−1.132		−4.769	
	(2.222)		(3.455)	
($N = 2$) Y/N		−1.659		0.150
		(10.542)		(10.446)
($N = 3$) Y/N		−11.313		−7.972
		(12.442)		(13.332)
($N = 4$) Y/N		−9.113		−7.192
		(17.904)		(17.347)
($N = 5$) Y/N		−17.223		−14.073
		(15.513)		(15.726)
($N \geq 6$) Y/N		8.074		12.122
		(15.339)		(20.606)
% fail both exams	−0.479	−0.490	−0.515	0.130
	(0.601)	(0.587)	(0.604)	(0.691)
Number of instructors	4.120***	4.237***	3.849***	3.156***
	(0.855)	(0.884)	(0.823)	(0.910)
Median weight of fleet cars (kg)	0.066***	0.067***	0.065***	0.045**
	(0.023)	(0.023)	(0.022)	(0.022)
School ≤ 100m to high school	58.871***	60.636***	58.416***	52.720**
	(22.178)	(22.682)	(21.682)	(25.331)
School distance to city center	1.578*	1.730*	1.674**	0.966
	(0.849)	(0.891)	(0.834)	(0.597)
% of population in urban parishes	−0.558***	−0.495***	−0.520***	−0.593***
	(0.173)	(0.173)	(0.172)	(0.204)
Per-capita income	0.003	0.001	−0.004	−0.033
	(0.040)	(0.041)	(0.040)	(0.036)
Number of car repair shops	0.535***	0.492***	0.601***	0.410
	(0.125)	(0.139)	(0.133)	(0.264)
Population (000)	−0.477*	−0.662***	−0.206	−0.553*
	(0.255)	(0.161)	(0.351)	(0.303)
Adjusted R^2	0.133	0.141	0.127	
1st stage partial R^2			0.689	
1st stage F statistic			16.614	
Observations	636	636	636	636

* p < 0.1, ** p < 0.05, *** p < 0.01. Municipality-level clustered standard errors in parentheses.
Note: For a list of suppressed controls and instruments, see footnote to Table 5.