Reference Pricing as a Deterrent to Entry: Evidence from the European Pharmaceutical Market*

Luca Maini† (Harvard University)

Fabio Pammolli‡ (Politecnico di Milano)

December 20, 2018

Abstract

This paper empirically and theoretically analyzes the impact of external reference pricing (ERP) on launch delays in the market for pharmaceutical products. Governments that implement ERP use prices in other countries as negotiation benchmarks to bring down the cost of prescription drugs. By doing so, they limit the ability of firms to price discriminate across countries and create an incentive to withhold drugs from countries with lower willingness to pay. Using data on pharmaceutical sales in European countries from 2002 to 2012, we document the presence of widespread launch delays across Europe — up to three years on average in Eastern Europe. To distinguish between strategic delays caused by ERP and delays that arise for other reasons, we develop a dynamic structural model of entry that allows for externalities in price, which we estimate using a novel moment inequality approach. We find that removing ERP would reduce delays in Eastern Europe by up to 14 months per drug. At the same time, strategic delays have a relatively small impact on firm revenue, so it would be theoretically feasible to compensate firms for their profit loss to remove strategic delays.

*We are extremely grateful to Ernst Berndt, David Cutler, Robin Lee, and Ariel Pakes. We would also like to thank Vincenzo Atella, Francesco Decarolis, Josh Feng, Joshua Gottlieb, Sonia Jaffe, Casey Mulligan, Myrto Kalouptsidi, Richard Manning, Tomas Philipson, Daniel Pollman, Mark Shepard, Ariel Stern, Elie Tamer, Pietro Tebaldi, Heidi Williams, Annetta Zhou, and participants of the Industrial Organization and Health Care Policy Seminars at Harvard, Applications of Economics Workshop at the University of Chicago, Bates-White Life Sciences Symposium, and EIEF Doctoral Workshop on Applied Microeconomics. Luca Maini thanks the Becker Friedman Institute for their financial support through the Health Economics Fellowship. IMS Health provided access to the data.

†lucamaini@fas.harvard.edu

‡fabio.pammolli@polimi.it

1 Introduction

In October of 2018, the US Department of Health and Human Services (HHS) published a policy proposal called International Pricing Index Model for Medicare Part B. If enacted, the policy would tie reimbursement for Medicare Part B drugs to international prices. The proposal has two main motivations. First and foremost, it is an attempt to rein in the rise of drug prices – a main concern for governments around the world. Second, it is an attempt to reduce the disparity in drug prices between the US and the rest of the world. The International Pricing Index Model is a form of external reference pricing (ERP), a class of policies defined by the WHO as the practice of benchmarking drug prices by using prices in foreign countries. Even though external reference pricing has never been adopted in the US, it is commonly used by governments abroad. ERP has an obvious appeal: it is simple, it guarantees that prices will be in line with other countries, and it can reduce spending. Firms however, worry about the limitations that ERP imposes on their ability to optimally price discriminate across countries with different willingness to pay.

The debate over the effects of reference pricing has important implications for policy. Using ERP carries virtually no negative repercussions for the home country, but it can impose an externality on foreign countries. Firms will have an incentive to delay entry in low-income countries whose prices are referenced by high-income countries whenever the externality imposed through reference pricing outweighs the revenue earned by expanding into additional markets. Drugs often enter in many markets several years after receiving marketing approval, so the welfare loss generated by ERP could be very large (Reich, 2000).

In this paper, we develop a structural model of entry that allows for externalities in price across markets and use it to provide an estimate of the impact of ERP on launch delays. In the model, firms maximize profits by choosing an optimal entry sequence, conditional on demand and price conditions in each country. To isolate the impact of reference pricing we exploit the fact that ERP operates through a very specific channel. On the margin, ERP generates delays when the change in expected profits from launching in an additional country is outweighed by the expected loss from the externality generated by reference pricing.

We estimate the model using data on sales of pharmaceutical products across all Member States of the European Economic Area (EEA). Most European countries adopt ERP among the criteria used to set prices. Moreover, the EEA includes countries with highly heterogeneous income, which creates the potential for strategic delays. Empirically, widespread launch delays occur in almost all European countries, though some of these delays are likely due to factors other than reference pricing. Before launching, governments and drug manufacturers engage in negotiations that can last several months, and countries can also delay the entry of products they consider unsafe. On average, relative to their initial marketing approval date, products are delayed by about 3 years in Eastern European countries, and 1 year elsewhere.

We begin our exercise by estimating demand and price in each country. We use a random utility nested logit model for demand, and estimate prices using a flexible parametric function that tries to capture the decision process of the government. The function predicts equilibrium prices as a combination of the reference price and an internal *government price* — which represents the price that would have been granted in the absence of reference pricing. The degree to

which the reference price affects the equilibrium is a country-specific parameter in the pricing model. Consistent with previous research we find that allowing this additional degree of heterogeneity is important, as many countries do not follow their stated ERP guidelines perfectly (Leopold et al., 2012a).

Our estimated demand and price primitives suggest that the externality generated by reference pricing is large enough to incentivize strategic delays. In simulations that compare the expected revenue of various entry sequences we find that firms earn higher revenue when delaying entry in some countries. In particular, delaying entry in at least one Eastern European country yields higher revenue for 70% of drugs. In roughly 20% of cases, withholding the product from *all* Eastern European countries would be preferable to launching everywhere at the same time. However, we find that virtually no drugs would earn higher revenue by delaying entry in any country outside Eastern Europe.

To quantify the impact of ERP we must isolate delays due to reference pricing from other sources of delays, such as the time needed to negotiate pricing and reimbursement conditions. We model delays that are not generated by ERP using a binary stochastic process. Every period, firms draw a shock for each country where they apply for entry. If it comes up negative, entry must be postponed until the next period, when a new shock is drawn.

The structure of the delay shocks accurately reflects the regulatory architecture faced by firm, but also presents a challenge to estimation because it implies that firm strategies are unobserved: situations when the firm applied and was delayed are observationally equivalent to situations in which the firm did not apply at all. This would not be an issue if we were able to solve the model. However, finding an exact solution to our model is not possible. There is no analytic solution, and the cardinality of the action space of the firm makes a numerical approach unfeasible (for a set of N countries over a T-period horizon, the firm can choose between T^N possible strategies).

In order to overcome this obstacle, we develop a novel moment inequality estimator. Our approach does not require us to identify the optimal strategy of the firm or compute the value function, though it can only provide bounds on the parameters of the model. Our inequalities rely on a revealed preference argument. We assume that firms are maximizing expected profits and compare the expected profits of the observed entry sequence to the counterfactual profits of playing a different strategy. These inequalities will not always hold for individual firms: the realization of the random delay shocks in the data might prevent the firm from achieving the desired entry sequence. However, these differences disappear in a large enough sample if we consider average payoffs across many firms. Our estimator generates moment conditions based on the payoff of the average drug and relies on a generalized version of the law of large numbers for non-identical, independently distributed random variables with finite mean and variance. In our empirical application we find that these moment conditions can only provide a lower bound on the parameter of interest. We calculate an upper bound by exploiting the fact that the approval date is the earliest time at which the firm could have sent an entry application.

Over the period from 2002 to 2012, our estimates imply that replacing ERP with a pricing mechanism that does not generate externalities in price across countries would reduce delays

in each Eastern European country by up to 63%, or 14 months per drug on average.[1] Several possible alternatives to ERP have been proposed in the policy literature, from transitioning to a centralized European cost-effectiveness evaluation system (Drummond, 2003), to implementing two-part pricing systems where products are supplied at cost and governments make transfers to firms in order to reach static and dynamic efficiency, to creating barriers preventing reference pricing and import-export of pharmaceutical products across countries (Towse et al., 2015). The exact policy would affect firm profits, but not the implications for strategic delays, so our counterfactual has broad external validity.

Using our estimates, we estimate that firms gain on average around €20 million by engaging in strategic delays, relative to what they would earn by playing a naive strategy that would see them apply for entry in all countries at the same time. While €20 million is obviously a large sum, it only represents a small percentage of the average lifetime earnings of drugs in the European market. This is a result of two factors. First, in the current equilibrium we estimate that prices in in Eastern European countries are not much lower than prices in Western European countries that use ERP aggressively. For example, the average price level across Eastern Europe is only slightly lower than the price level in Italy and Spain.[2] Second, the structure of the delay shocks mitigates the impact of ERP, as we estimate that countries that grant lower prices, also tend to be slower in reviewing pricing and reimbursement applications. The fact that firms gain only a little by strategically responding to ERP suggests that it may be relatively inexpensive to reduce delays by offering a lump-sum subsidy in exchange for sending entry applications to all countries simultaneously.

Our analysis has some limitations. The complexity of the problem we consider forces us to make several simplifying assumptions, which are necessary to perform the empirical analysis, but not necessarily desirable. First, we assume that firms act as single-agents. Even though the firms we consider are monopolists with regards to the specific molecule they produce, virtually all therapeutic classes we consider contain at least a few different molecules, which are presumably substitutable with one another to some degree. Second, we assume that there is no structural error in either our demand or price model. This assumption is necessary to build the moment inequalities given that firm strategies are unobserved. Finally, we do not explicitly model the government's choice of a reference pricing function, opting instead to treat it as an exogenous feature. We discuss these limitations more in depth in the relevant sections of the paper and in the conclusion.

Our paper contributes to four main strands of economic literature. First, it belongs to a growing body of work, both empirical and theoretical, studying how price regulation affects access to pharmaceutical products. The empirical side of this literature usually analyzes the impact of government policy on launches using a reduced-form framework (Danzon et al., 2005; Danzon and Epstein, 2012; Kyle, 2007; Kyle and Qian, 2013; Cockburn et al., 2016). Two

[1] By distinguishing between Eastern Europe and Western Europe we do not mean to associate any value to the geographic location of these sets of countries. Rather, we draw this demarcation for convenience. Countries in Eastern Europe share certain traits that make their bundling convenient for our purpose: they have lower income (and prices), and smaller market size than virtually all countries in Western Europe.

[2] This is almost certainly an equilibrium result. Young et al. (2017) find that prices in Eastern European countries are higher-than-expected when compared to income levels, and argue that if reference pricing were removed, firms would almost certainly be willing to grant lower prices in Eastern Europe.

notable exceptions are Duso et al. (2014), which examines the welfare impact of parallel trade in Germany, and Dubois et al. (2018), which assesses the effect of a hypothetical US reference pricing policy, but does not internalize the impact of the policy on launch delays.[3] On the theory side, this literature has focused on simulating the impact of reference pricing on firm strategy (e.g. Borja, 2014; Toumi et al., 2013; Stargardt and Schreyögg, 2006; Houy and Jelovac, 2015), or establishing conditions under which regulation that limits price discrimination is beneficial or harmful to welfare (e.g. Birg, 2016; Brekke et al., 2007, 2015, 2016; Matteucci and Reverberi, 2017). Our contribution is that we explicitly model the impact of reference pricing on firm incentives and develop an estimation strategy to isolate the effect of this policy on launch delays.

Second, our paper is related to a series of studies on the impact of regulation that links prices to endogenous market benchmarks. For example, both Medicare Part B and Medicaid tie drug reimbursements to the average of reported private market prices. Duggan and Scott Morton (2006) show that in the case of Medicaid this regulation creates a distortion that leads to higher prices in the private market. Another set of policies with a similar effect are so-called "price-linked" subsidies, i.e. subsidies that are linked to market prices. Jaffe and Shepard (2017) and Decarolis (2015) show that these types of subsidies can distort premiums in health exchanges and Medicare Part D respectively. More generally, price externalities across firms have been detected in the absence of government intervention. Grennan (2013) and Grennan and Swanson (2016) show that knowing how much rival hospitals paid for medical devices can affect future prices. Our paper shows that if pricing strategies are constrained, firms can also respond along different margins (i.e. by manipulating their entry strategy).

Third, we contribute to the vast empirical Industrial Organization literature on entry models, which originated with Bresnahan and Reiss (1991) (for an overview of this literature see Berry and Reiss, 2007). Most papers in this literature use stochastic fixed costs of entry (e.g. Seim, 2006). Since these costs are less relevant in our setting, we take a different approach and replace them with stochastic delay shocks. These shocks have different implications for estimation: they do not directly affect the profit of the firm, but rather impose stochastic constraints on the action space. Our paper derives conditions which these parameters under what conditions these parameters can be identified.

Our final contribution is to the literature on partial identification started by Manski (2003). We develop a novel approach to deal with the challenges introduced by delay shocks. The empirical literature on partial identification is growing and includes several papers (Dickstein and Morales, 2018; Eizenberg, 2014; Holmes, 2011; Illanes, 2016; Katz, 2007; Morales et al., 2017; Pakes et al., 2015). Our approach is closest to that of Holmes (2011) and Morales et al. (2017), but differs in the way that identification is obtained. While their approach identifies the set of parameter values for which the firm's observed strategy is optimal, our approach identifies the set of parameters consistent with the revenue earned by the firm in the data.

The rest of the paper proceeds as follows. Section 2 introduces the institutional environment of the European pharmaceutical market. Section 3 describes the data and discusses pre-

[3] Another methodologically related paper is Chaudhuri et al. (2006), which uses structural techniques to estimate the impact of patent policy on patient welfare in the Indian market for quinolones.

liminary evidence that supports the hypothesis that firms are delaying launches in countries with low prices in order to avoid the impact of external reference pricing. We present our theoretical model of entry in Section 4. The estimation is then divided in two parts. We present our empirical model and estimation results for demand and price in Section 5, while Section 6 contains the dynamic analysis. We discuss the implications of our results for counterfactuals and policy analysis in Section 7. Finally, in Section 8 we provide some concluding remarks, a discussion of the paper's limitations, and a roadmap for future research.

2 Overview of the European Pharmaceutical Market

2.1 *Marketing Approval and Price Regulation of Pharmaceutical Products in Europe*

New drugs can only be sold after being reviewed for efficacy and safety. The European Medicines Agency (EMA) oversees this process in the European Economic Area. While marketing approval for new drugs is generally granted by a regulatory authority whose jurisdiction is limited to one country (i.e. the FDA in the United States, or the PMDA in Japan), member states of the European Economic Area have been relying on a shared approval process since 1995 – the year the EMA was founded.[4] Though national drug agencies still exist, their effort is now organized and regulated by the EMA.

Pharmaceutical companies seeking approval for their products can choose between three possible procedures. The *centralized procedure* is administered by the EMA itself, and grants automatic approval in all EEA Member States. It is available to all drugs, and compulsory for certain classes of drugs, including biologics.[5] Drugs for which the centralized procedure is not mandatory can also go through two additional channels. If the drug already has a marketing authorization from any EEA member state, the firm can use the *mutual recognition procedure* to extend it to any other member state using a fast-track procedure taking no longer than 90 days (European Parliament, 2001).[6] The other alternative is the *decentralized procedure*. In this case, the firm submits an application to multiple countries at the same time and designs one as the Reference Member State in charge of reviewing it (European Parliament, 2004).

The centralized marketing approval process ensures that the cost of seeking additional marketing approvals all but disappears as soon as firms receive marketing approval from any country in Europe (or from the EMA). This eliminates one of the most common explanations for launch delays. However, firms may still incur into delays because individual countries retain the ability to regulate prices independently from one another.

Most European countries provide some form of single-payer coverage, meaning the government bears the vast majority of prescription drug costs. Even where the government does

[4]The European Economic Area consists of all Member States of the European Union, plus Norway, Iceland, and Liechtenstein. Switzerland (also in our data), is a member of the European Free Trade Area, but not of the EEA. However, it has a series of bilateral trade agreements that allow it to take part in the common European market.

[5]The full list of drugs that must receive approval by the EMA includes: human medicines containing a new active substance to treat HIV or AIDS, cancer, diabetes, neurodegenerative diseases, auto-immune and other immune dysfunctions, and viral diseases; medicines derived from biotechnology processes; advanced-therapy medicines, such as gene-therapy, somatic cell-therapy or tissue-engineered medicines; and medicines seeking orphan designation.

[6]Other countries can to refuse the extension by claiming that doing so would create significant risks for public health, although this does not happen very frequently.

not directly insure patients (e.g. Germany, Netherlands, Switzerland), it still negotiates a price cap with manufacturers. Thus, all governments impose strict restrictions on drug prices, with the primary goal of controlling spending.

Pricing restrictions typically only apply to drugs that are paid for by the government through the public health insurance system. However, since European citizens overwhelmingly access health care through government-funded programs, the exclusion of a product from public formularies results in its de-facto exclusion from the national market (European Commission, 2012).[7] The ability to effectively deny entry provides governments with the necessary leverage to demand lower prices.

Firms petition for reimbursement status by submitting pricing and reimbursement applications to each government. The time required to review an application and negotiate a price can vary significantly across countries. In theory, Directive 89/105/EEC (informally known as the Transparency Directive) states that governments can take no longer than 180 days to review a pricing and reimbursement application (Council of European Communities, 1988). In practice however, this limit is often surpassed, both because of enforceability issues, and because governments can stop the clock by asking for additional information. Data on turnaround times for applications is scarce, but survey evidence from the late 1990s indicates that the average varies substantially, from 0 days in UK and Germany, to over two years in Belgium (OECD 2008; PICTF, 2006).

The requirements of pricing and reimbursement applications vary across countries, though firms must generally include both a clinical dossier detailing the medical benefits of the drug, as well as an economic report with projected sales and a proposed price. The government then uses this information as inputs into the pricing decision. In theory, the government strives to set prices that reflect the value of each drug and reward the firm's innovative efforts, while at the same time keeping spending under control. In practice, estimating the value of a drug is a complicated and costly exercise. Most countries use ERP as a way to set prices that are approximately consistent with what other governments are paying.

It is important to stress however, that ERP is not the only (or even the main) policy instrument used by countries in setting prices. Countries often rely on a variety of other methods, including, but not limited to Health Technology Assessments, internal reference pricing (which links prices of molecules within a pre-specified equivalence class), and price freezes/cuts. These policies will be especially important in situations when a reference price cannot be observed, or when the reference price is higher than what the government is able (or willing) to pay.

Finally, governments ultimately care about overall pharmaceutical spending rather than prices. Price-volume agreements that protect spending such as clawback policies are common, meaning that the outcome of the negotiation between the government and the firm is usually not a set list price, but rather a price schedule that depends on volume (Carone et al., 2012). Prices and volumes will also be correlated if the government is willing to provide more favorable coverage to firms that give better pricing conditions.

[7] The only exceptions to this rule tend to be generic drugs, whose low price tag makes them a cost-effective option even in the absence of government coverage.

2.2 *Overview of External Reference Pricing*

In 2012 all European countries indicated ERP as one of the criteria used in setting prices except Denmark, Germany, Sweden, and the UK.[8] Both the pharmaceutical industry and policymakers have acknowledged the externality generated by ERP and its role in producing launch delays for new pharmaceutical products (EFPIA, 2014; Carone et al., 2012). However, governments remain reluctant to abandon the policy, because of the savings they claim it generates.

The two most important aspects of ERP policies are the reference basket (i.e. the basket of countries whose prices are sampled), and the formula used to compute the reference price. For both, there is significant variation across countries.[9] Some governments (e.g. Austria, Belgium, Finland, Hungary, and Poland) require firms to submit prices from all other countries in the European Union. Others only reference similar countries, both in terms of geographical proximity, size, and income level — for example, Estonia references Hungary, Latvia, and Lithuania, while France references Germany, Italy, Spain, and the UK. In terms of the reference formula, most countries use the average across the reference basket, but a few (e.g. Latvia, Poland, and Romania) use the lowest price, while others still use slight variations: Bulgaria, Greece, and Norway use the average of the three lowest prices in the basket. Figure 1 offers an overview of cross-country variation in reference baskets and formulas.[10]

The stringency with which each country adheres to their stated ERP guidelines may vary across countries. Some governments state that ERP is only used to "inform" the pricing decision, meaning that we might expect prices to be affected by ERP but not necessarily to be perfectly aligned with the reference pricing benchmark. In other instances, governments may push for prices that are below the benchmark if they expect to sell higher volumes than the referenced countries.[11] Countries whose governments claim to only use ERP informally include Belgium, Finland, France, Italy, Poland, and Spain.

We use the reference basket and formulas to estimate reference prices in our model, but we exclude a few additional characteristics of ERP policies that can also vary across countries. We briefly list them here for completeness. First, countries update the reference prices with varying frequency, from as little as every 6 months (e.g. Greece, and Slovenia), to as many as 60 (Finland). Second, countries can use raw ex-factory prices, or apply a PPP adjustment (all Scandinavian countries do so). Third, not all countries apply ERP to the same set of drugs. Most countries apply ERP only to drugs that are reimbursed through the national health in-

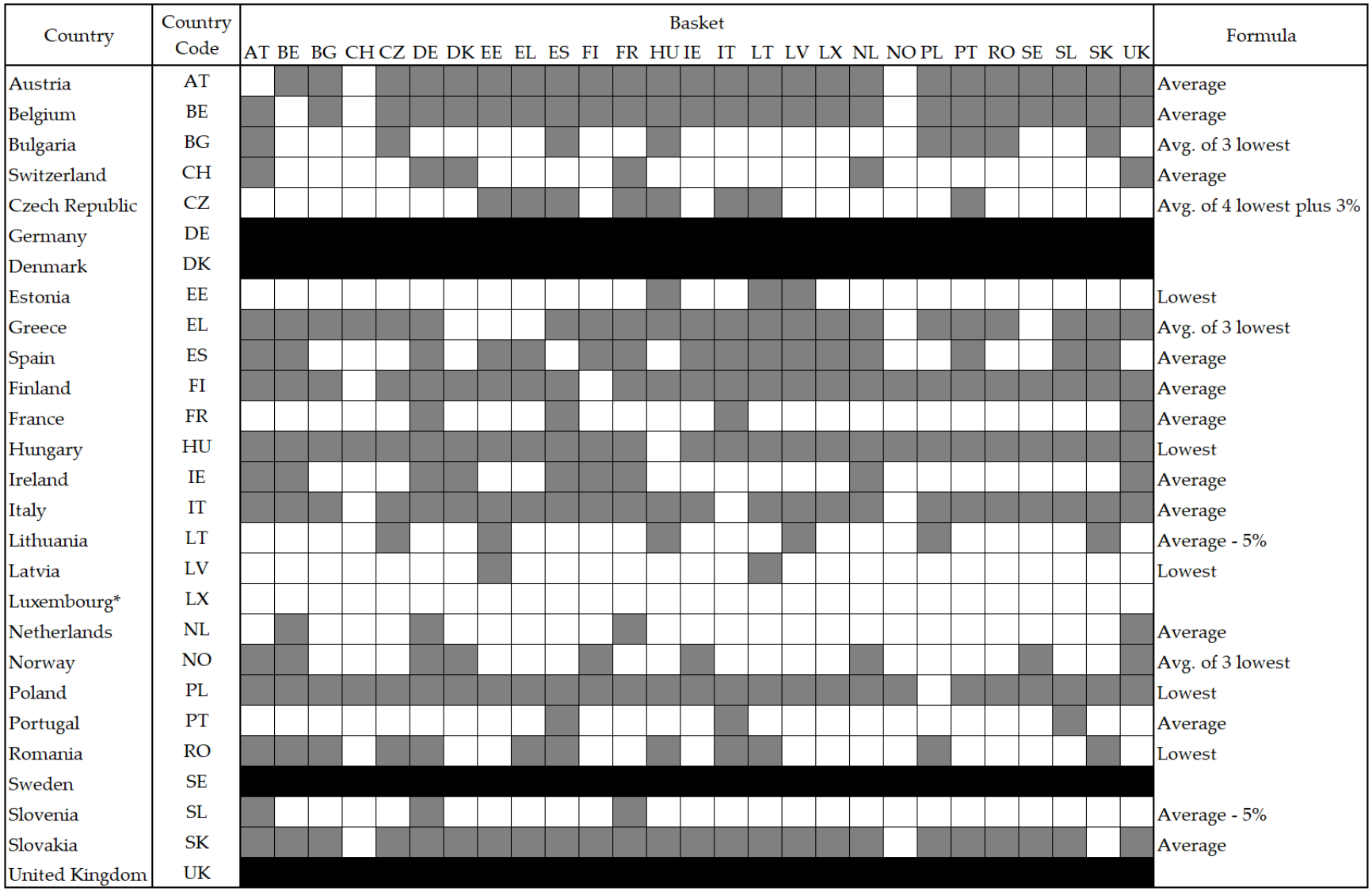
[8] Since then, Denmark and Germany have adopted ERP as well.

[9] Contrary to cross-country variation, over time variation in ERP policies is much more scarce. The biggest policy change happened in 2010 when Greece switched its reference function in an effort to lower prices. Other changes are tied to the entry of new Member States in the European Union or in the Eurozone. These occurred in Austria, Belgium, Finland, and Italy (all of whom reference EU member states); and Spain (which references countries in the Eurozone). Finally, Portugal, Hungary, and Poland made small adjustments to their reference baskets at various points. Please refer to the Online Appendix for more details.

[10] The table shown in the paper represents a snapshot of reference baskets and formulas in 2012. It was built by combining several published sources (Carone et al., 2012; European Federation of Pharmaceutical Industries and Associations, 2014; Kanavos et al., 2011; Leopold et al., 2012b; Wilsdon et al., 2013) with unpublished IMS reports. For the analysis used in the paper we generated yearly tables to capture some small changes that happened in a few countries with regard to their reference basket and the formula used.

[11] According to several informal conversations of the authors with industry insiders, these sort of arguments appear to be popular in countries such as Italy, where the prices of smaller markets are included in the reference function.

Figure 1: REFERENCE PRICE BASKETS AND FORMULAS FOR EEA COUNTRIES



* Luxembourg only references the drug's country of origin.

This figure shows the reference baskets and formulas for all European countries. Each row shows the basket of the country in the first column. Grayed-out squares indicate which countries belong to the basket. The sources used to compile the Figure are described in footnote 10 on the preceding page.

surance system, but some apply it to all new innovative drugs (e.g. France), and others to all drugs, regardless of reimbursement status (e.g. Greece).

2.3 Data and Summary Statistics

The main source of data for the empirical analysis is the MIDAS database maintained by IQVIA (formerly IMS Health), a global information company specializing in the health care sector. The data covers sales of all pharmaceutical products for European countries from 2002 to 2012.[12] It consists of a quarterly panel of volume and revenue sales divided by country. Products are defined by a combination of molecule, firm, product name, form, strength, and package. IQVIA collects this information by surveying pharmacies and hospitals.

To the best of our knowledge, this database represents the most comprehensive source of data on sales in the European pharmaceutical market. Nonetheless, it has a few important limitations, which we discuss below.

First, the data does not provide any information on the approval dates of drugs. We collect approval dates for all EMA-approved medications from the EMA's website, and the approval date of all mutual recognition applications from an internet database maintained by the Heads of Medicines Agencies (HMA).[13]

Second, IMS reports ex-factory revenue sales, which do not usually incorporate rebates and discounts that are sometimes granted by individual payers. While in the US estimates of discounts for brand drugs oscillate between 20-40% during the period we consider (see Congressional Budget Office, 2005; Aitken et al., 2016), discounts tend to be much lower in Europe (Danzon, 2003; Danzon and Furukawa, 2006). According to industry insiders, average rebates for patent-protected brand drugs are rarely above 10%.[14] Unfortunately, there is currently no available data on pharmaceutical rebates in Europe, so in our paper we simply use the prices implied by the IMS data.[15]

Third, the data contains some missing information for certain countries and years. Because of the externalities generated by reference pricing, missing data points can have an impact on non-missing observations as well. To minimize the impact of missing data we resort to imputation using a variety of techniques.[16]

We integrate the IMS data with a few additional sources. On top of the aforementioned EMA and HMA data on approval dates, we collect GDP and population data from Eurostat, as well as data on the incidence of diseases in each European country from the Global Burden of Disease Study. We use that information to build the market size variable for the demand

[12] We are missing data entirely for Cyprus, Iceland, Lichtenstein, and Malta. A few other countries have partially missing data. See the Online Appendix for more details.

[13] Both datasets are publicly available. The HMA is a network of the heads of the European national authorities in charge of the regulation of medicinal products for human and veterinary use in the European Economic Area. The data can be found at http://mri.cts-mrp.eu/Human

[14] Author's own estimation based on several conversations with industry insiders.

[15] As an aside, contracts that combine list prices with hidden discounts could circumvent ERP and restore the ability to use price discrimination. However, this does not seem to be the case even though these contracts are theoretically available. One possible reason why rebates are underutilized in Europe is that governments have higher transparency requirements than private firms, due to the necessity to account for their spending. Hence hiding discounts might be more difficult. Another possibility is that firms know that if they were to avoid ERP by using large discounts, government might find ways to force them to reveal those discounts.

[16] See the Online Appendix for more details.

estimation. We also use quarterly exchange rates from the European Central Bank to convert sales data from countries that use currencies other than the Euro.

We observe around 6,000 molecules and 3,000 firms in our data. Most of these molecules and firms are old off-patent and generic products with negligible yearly sales, and many are only available in one or two countries. We are interested in new, on-patent products whose potential market spans multiple European countries. Hence, we select a subsample of drugs that satisfy the following three criteria: the drug was first launched on or after January 1st, 1995; it had at least one new launch in a European country on or after January 1st, 2002; and it was either approved by the EMA using the centralized procedure, successfully completed at least one Mutual Recognition Application between 1995 and 2012, or is a patent-protected brand drug sold in at least 10 countries by 2012.

Our final selection comprises 481 drugs (we define a drug as the combination of a molecule, a firm, and a therapeutic class).[17] Most of the products we select received approval from the EMA through the centralized procedure or applied for mutual recognition. We also include a few drugs that we were not able to match with the EMA and HMA data on approval dates, but that we observe being sold in many European countries.[18] Unsurprisingly, drugs in our main sample experience much greater sales and diffusion relative to the average drug in the data (see Table 1). The median drug in our main sample is available in 22 countries (by the end of 2012), and collects yearly sales of €38.6 million across all European markets. For comparison, the median product in the full sample is sold in only 1 country, and earns less than €100,000 every year.

We also select a subsample of drugs within our main sample whose patent expired prior to December 31st, 2012. This smaller group is used in the dynamic analysis, when our methodology requires that we are able to compute the overall expected payoff of a drug until the time its patent expires.[19] This smaller sample consists of 87 drugs and has similar characteristics relative to the main sample of drugs in terms of sales and diffusion across European countries. The only main difference is in how these drugs were approved. Most of the drugs in our main sample were approved by the EMA using the centralized procedure. However, a majority of the drugs in the dynamic sample chose the Mutual Recognition Procedure.

Relative to the approval date, virtually all countries experience meaningful launch delays on average, though there is substantial heterogeneity across countries. Figure 2 reports delays by country using the full unbalanced sample of 481 main products.[20] The graph plots average

[17] All molecules in our main sample are sold by a single firm, but can sometimes be available in different therapeutic classes. In these cases, IMS reports separate observations for each therapeutic class. We keep this distinction since our demand estimation relies on therapeutic classes to define markets.

[18] For these products we impute the European approval date as the date of the fifth launch. We do not use the first launch because in many instances products that apply through the mutual recognition procedure start selling in the reference member state a year or two in advance relative to the rest of Europe.

[19] Since patent expiration dates can vary slightly across countries, we set period T as the latest expiration date among those of France, Italy, and Spain. Patents generally expire roughly at the same time in most countries, since they are administered by the European Patent Office. However, some countries can choose to grant extensions to individual patents. In our data we also occasionally observe earlier than expected patent expiration dates for some Eastern European countries. We choose these three countries because they are the three largest markets that use ERP. Therefore, when their patent protection expires, the strategic incentives to delay launches should become close to zero. Empirically, we observe only a total of 11 country launches occurring after period T, which suggests that our approximation is accurate.

[20] Note that this means that the sample is truncated. For a similar look at delays in a balanced (but smaller)

Table 1: SUMMARY STATISTICS

		Full Sample	Main Sample	Dynamic Sample
# therapeutic classes		241	109	44
# firms		2,944	168	47
# molecules		6,354	475	86
Class-firm-mol combinations		55,134	481	87
Class-firm combinations			375	84
Diffusion	mean	2.1	20.1	21.3
	median	1	22	24
Yearly sales	mean	€3,547,665	€115,427,475	€121,977,298
	median	€92,489	€39,214,276	€46,340,438
Approval Method	EMA		312	24
	MRP		127	46
	Other		42	17

This table reports summary statistics for the IMS MIDAS database. The full sample includes all prescription drugs in the data. The main sample includes prescription drugs that satisfy the criteria laid out in section 2.3. The dynamic sample includes a subset of main sample drugs whose patent expired by December 31st, 2012. See Section 2.1 for an overview of the marketing approval methods. Yearly sales refers to the entire EEA territory.

launch delay in months, conditional on entry and calculated starting from the approval date of each product. The average delay ranges from as little as 3 months in the Netherlands, to as much as 3 and a half years in Romania.

Some of this heterogeneity is explained by income: low-income Eastern European countries experience on average 2 additional years of delay relative to the rest of Europe. In Figure 3 we document a strong inverse relationship between price level and delays, and show that prices in countries that use ERP fall over time relative to prices in countries that do not use ERP. To obtain the data used to generate the Figure we calculated price level in each country using a fixed effect regression, and plotted the coefficients against average delays from Figure 3.[21] There are three main takeaways from this figure. First, price levels and delays are negatively correlated, as we would expect if firms delayed entry in countries that cannot afford to pay high prices. Second, countries with a large population tend to experience shorter delays relative to the benchmark set by the best-fit line, while countries with low population tend to experience longer delays. This is also consistent with an explanation for delays that is based on reference pricing: firms will be more willing to surrender a low price in large countries since the larger market size will make up for the potential loss from the externality on price. Third, the effect of market size is asymmetric: low-income countries are consistently penalized for market size, but small high-income countries are equally likely to fall above or below the best-fit line. This is once again consistent with a strategic reaction to ERP: small market countries

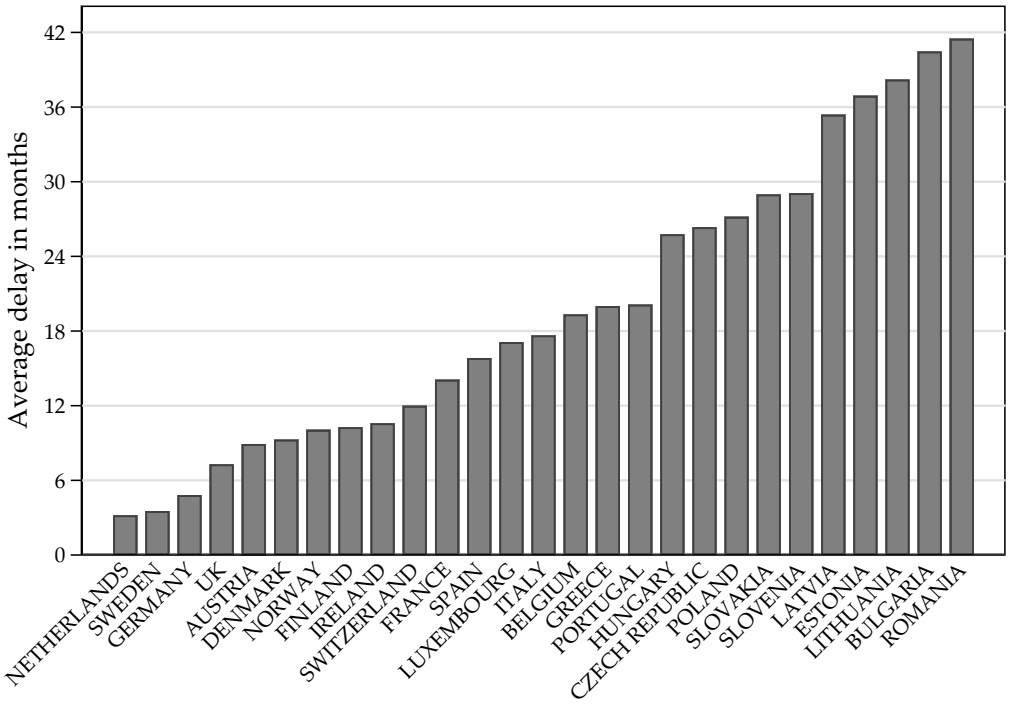
sample, see the Online Appendix.

[21] Average delay comes from Figure 2. Price level variable is calculated as a country-specific fixed effect γ_j in a regression of log price on various fixed effects:

$$\ln\left(p_{ijt}\right) = \theta_i + \gamma_j + \delta_t + \varepsilon_{ijt}$$

In the regression, i indexes drugs, j indexes countries, and t indexes years.

Figure 2: AVERAGE DELAY FROM MARKETING APPROVAL BY COUNTRY



This graph shows the average launch delay in months conditional on entry. Delays are measured starting from the approval date and are calculated using all 481 drugs in the main sample. See Table 1 on the previous page for summary statistics of this group of drugs, and section 2.3 on page 10 for a description of how this sample was created.

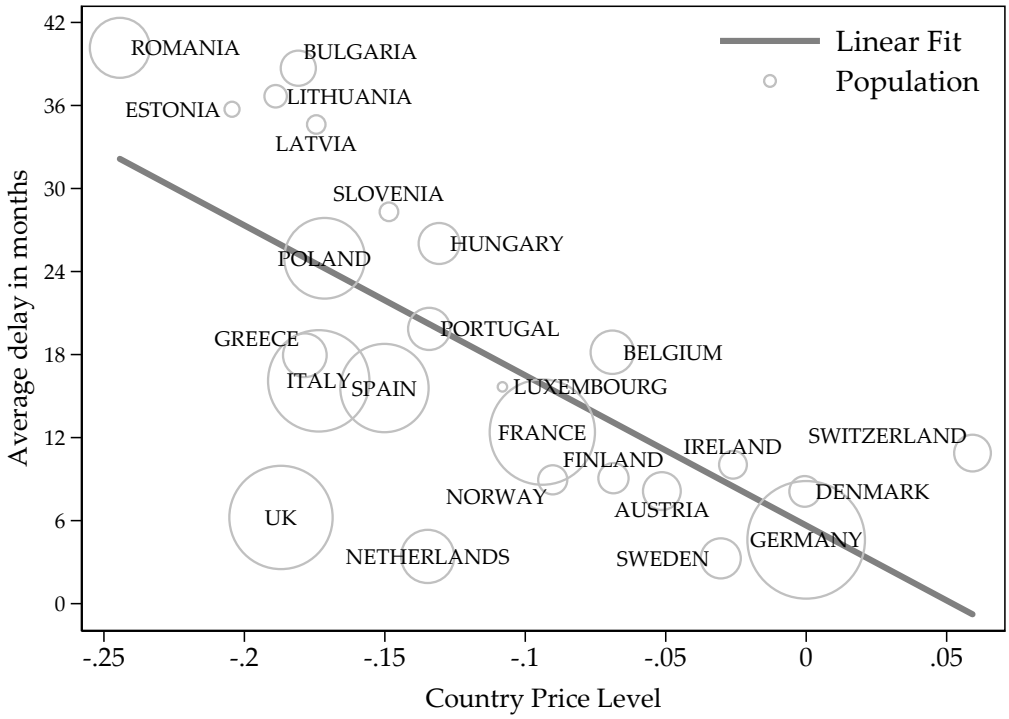
where prices are high do not generate a negative effect on prices, and therefore should not be penalized.

3 PRELIMINARY EVIDENCE OF THE IMPACT OF ERP

In this section we present suggestive evidence that delay patterns observed in Europe are inconsistent with models that do not incorporate reference pricing. The presence of delays does not represent, in and of itself, enough evidence that firms are responding strategically to reference pricing. Various alternative models would justify delays: they could be caused by fixed costs of entry, capacity constraints, or because firms can only send a limited number of entry applications to each country. We show that the comparative statics that these models predict are inconsistent with the delay patterns we observe in the data. Instead, our evidence suggests that they are driven by reference pricing.

We begin by introducing a stylized model of entry. Suppose a monopolistic firm has a license to sell a new drug in two countries. The drug has a lifetime of two periods, after which a generic enters, and profits fall to zero. For simplicity, assume that prices are set through an exogenous mechanism, so the firm's only choice variable is the launch sequence, which we denote as (s_1, s_2), where s_j is the period in which the product is launched. We also assume that demand and prices are constant over time, and that there are no costs of production. Denote demand and price in country j as q_j and p_j respectively and assume WLOG that $p_1 > p_2$.

Figure 3: CORRELATION BETWEEN PRICE LEVEL AND AVERAGE DELAYS BY COUNTRY



This figure plots the relationship between price level average delay in months. Price level is calculated as the country fixed effect from a regression of log price on drug, year, and country fixed effects. The coefficient for Germany was normalized to 0 in this regression. Average delays are the same as in Figure 2. The plot is weighted by population (data from 2012).

Within this framework we consider four possible motivations for delays: reference pricing, fixed costs of entry, capacity constraints, and limits to the number of launches that can be successfully completed in each period.

External reference pricing. In this scenario, assume that differential pricing can only be sustained for one period. After that, if the drug is available everywhere, governments take notice of each other's prices and demand the lowest one. In this scenario, the optimal strategy can either be to launch immediately in both countries or wait until the second period before launching in the country with a lower price.[22] Let the profits of a given launch sequence (s_1, s_2) be expressed as $\pi(s_1, s_2)$. Then

$$\begin{aligned} \pi(1,2) &= p_1 q_1 + (p_1 q_1 + p_2 q_2) \\ \pi(1,1) &= (p_1 q_1 + p_2 q_2) + (q_1 + q_2) p_2 \end{aligned}$$

Hence, a delay will occur if and only if

$$\pi(1,1) > \pi(1,2) \iff q_1 (p_1 - p_2) > q_2 p_2 \tag{1}$$

[22] Any other possible strategy is clearly suboptimal: launching in the country with a high price in the first period is better than both not launching at all, and launching in the country with a lower willingness to pay. Moreover, since prices adjust after one period, there are no downsides to launching everywhere in the second and last period.

The LHS of this equation represents the loss caused by reference pricing in the second period, while the RHS represents the additional sales from anticipating entry in country 2. Notice that since the loss depends on the difference in price between country 1 and country 2, delays in country 2 should be inversely correlated with prices even after controlling for revenue.

Fixed costs of entry. In this scenario, assume that in order to enter in a country the firm must pay a stochastic fixed cost of entry $\xi_{jt} \sim F_{j\xi}\left(\theta_{j\xi}\right)$. We can treat the entry problem in each country separately. In period 1, the firm decides whether to delay or not based on

$$\max\left\{2p_jq_j - \xi_{j1}; p_jq_j - \mathbb{E}\left[\xi_{j2}\right]\right\}$$

In particular, there will be a delay in country j in period 1 if and only if

$$\xi_{j1} > p_jq_j - \mathbb{E}\left[\xi_{j2}\right] \tag{2}$$

According to this model, the probability of delay should respond to revenue, but should not depend on price once revenue is accounted for.

Capacity constraints. In this scenario, assume that the firm has unlimited capacity in period 2, but can only produce a fixed amount $\bar{q} < q_1 + q_2$ in period 1. In this scenario, the firm would sell first in the country with a higher price, that is, country 1. Then, if $\bar{q} > q_1$ it would sell the remaining units in country 2. Delays in country 2 arise if $\bar{q} < q_1$. This model too, predicts that delays should be inversely correlated with price even after accounting for revenue.

Limited number of applications. In this scenario, we assume that firms can launch in at most one country in each period. In the first period, the firm will choose to launch in country j if and only if

$$p_jq_j = \max_{k\in\{1,2\}}\left\{p_kq_k\right\} \tag{3}$$

Like in the model with fixed costs of entry, prices shouldn't matter once revenue is accounted for in this model.

While delays can arise in all four variations of the model, each scenario predicts different delay patterns. In the scenarios that justify delays using fixed costs of entry and limits to the number of applications, revenue is the only variable that affects the decision of the firm to delay entry. Conversely, the variations that explain delays using external reference pricing and capacity constraints suggest that prices can affect delays even after controlling for revenue. Under the ERP regime, launching in a small country with a higher price may be more desirable than launching in a country with higher expected revenue, but a lower unit price. More generally, holding revenue constant, delays are more likely in countries with lower prices because a lower price generates a bigger externality. In the scenario with capacity constraints prices matter because in the absence of other costs, the firm would like to sell their stock wherever they can get a higher unit price for it.

In the data, we find that delays are inversely correlated with prices, even after controlling for revenue (Table 2). We test this hypothesis by regressing delays on revenue and price:

Table 2: IMPACT OF PRICE AND REVENUE ON DELAYS

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln\left(\text{Yearly Rev}_{ij}\right)$	-2.716^{***}	-3.663^{***}	-2.756^{***}	-3.769^{***}	-2.756^{***}	-3.757^{***}
	(-0.091)	(-0.097)	(-0.092)	(-0.097)	(-0.092)	(-0.096)
$\ln\left(\text{avg}\left(P_{ijt}\right)\right)$	-0.172^{**}	-1.509^{***}				
	(-0.087)	(-0.472)				
$\ln\left(P_{ijt_0}\right)$			-0.173^{**}	-6.015^{***}		
			(-0.088)	(-0.727)		
$\ln\left(\max\left\{P_{ijt}\right\}\right)$					-0.164^{*}	-8.771^{***}
					(-0.087)	(-0.719)
Drug F.E.	N	Y	N	Y	N	Y
R^2	0.09	0.38	0.09	0.38	0.09	0.39

$$\text{Delay}_{ij} = \alpha_i + \ln\left(\text{Yearly Rev}_{ij}\right) + \ln\left(p_{ij}\right) + \varepsilon_{ij}$$

To address the issue that revenue will mechanically be lower in countries where entry is delayed, we use average yearly revenue. With price, the main concern is that what we observe in the data is a combination of a country's true underlying price, and the influence of reference pricing. We test three different measures of price: average price, price at launch, and maximum price. Each has some advantages and disadvantages. Average price is more robust to year-to-year fluctuations but will also suffer the most from the impact of reference pricing. Price at launch and maximum price are less likely to be influenced by reference pricing but are also potentially more noisy.

Our results show that while the exact measure of price matters for the magnitude of the correlation, all three measure are significantly and negatively correlated with delays. This result suggests that the models with fixed costs of entry and limits to the number of applications are not capturing the full story behind delays.

Distinguishing between reference price and capacity constraints is harder just by looking at the correlation of price and delays because both models predict that delays are more likely to occur in countries with lower prices even after controlling for revenue. The ideal test of the capacity constraint model would require data on capacity, which is usually hard to come by. Instead, we use total output as a proxy for capacity and exploit the fact that some drugs have declining sales towards the end of their life-cycle. In the data, we can see the year in which the firm reached peak output and calculate the fraction of countries where the product is not yet available after output has already started declining. We find that when firms reach maximum output, the product still has not been introduced in ~20% of countries (Figure 4). This suggests that capacity constraints cannot entirely explain delays.

Finally, we run an indirect test of the model using reference pricing by looking at how prices change across countries that use ERP and countries that don't. To do so, we run the following regression:

$$\ln\left(p_{ijt}\right) = \theta_i + \gamma_{ja} + \delta_t + \varepsilon_{ijt}$$

Figure 4: MISSING COUNTRIES AT PEAK VOLUME OUTPUT

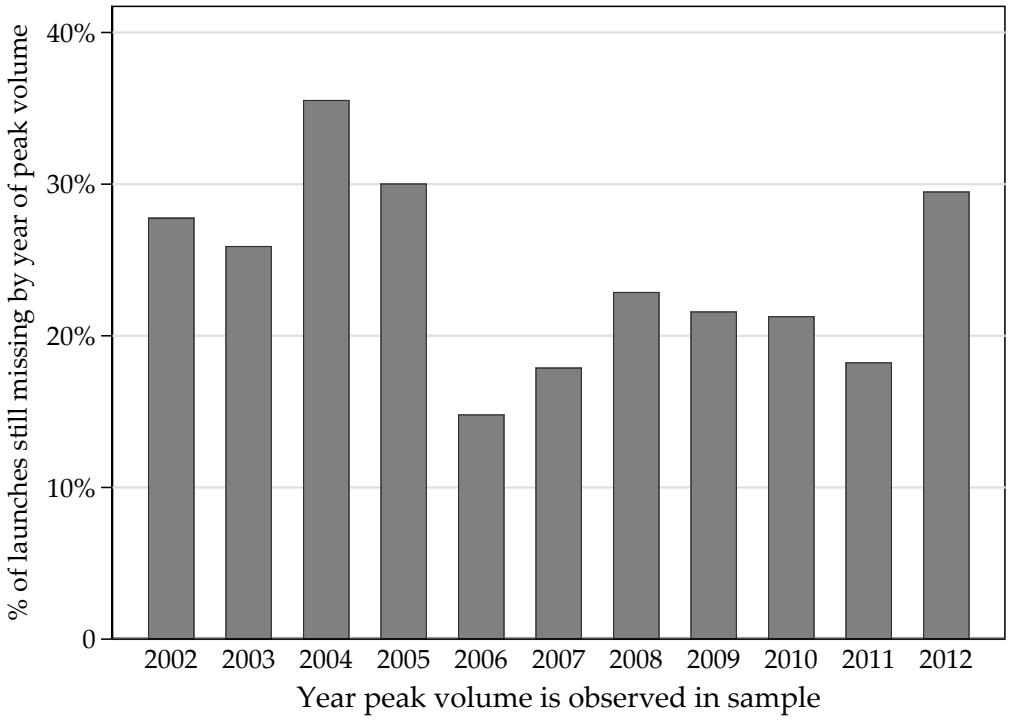
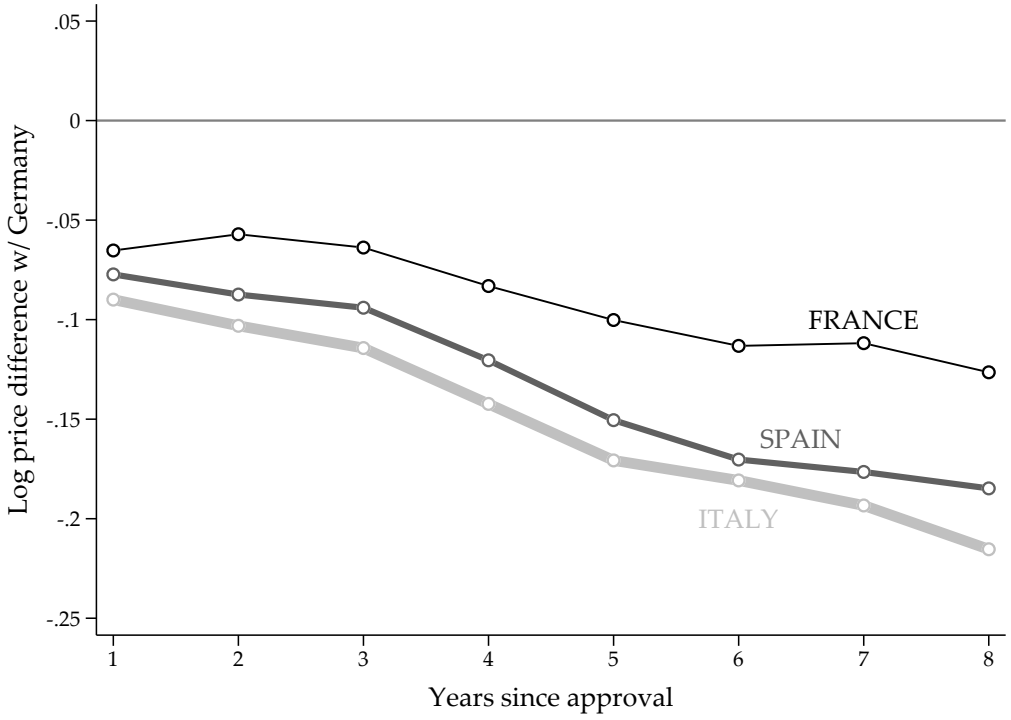


Figure 5: CHANGE IN PRICES OVER TIME AND ACROSS COUNTRIES



where γ_{ja} is a fixed effect for country and drug age, measured in years starting from the approval year. We also include drug fixed effects θ_i and year fixed effects δ_t. We focus on the four largest countries in the Eurozone: France, Germany, Italy, and Spain. Germany does not use ERP, while the three other countries do. To check whether prices diverge over time we plot the difference between the γ_{ja} coefficients for Germany and those of the the three other countries (Figure 5). The results show that relative prices in France, Spain and Italy fall over time relative to Germany. This is consistent with the additional downward pressure that we would expect to see through the external reference pricing channel: as the product is launched in more countries, prices fall wherever ERP is used relative to countries where ERP is not used.

4 A Dynamic Model of Entry with Externalities in Price

In this section we formalize the stylized model introduced in the previous section and extend it to obtain a statistical model for structural estimation.

A pharmaceutical firm l owns a set $\mathcal{I}_l$ of patent-protected molecules (indexed by i) with a marketing authorization for sale in a finite set $\mathcal{N}_i \subseteq \mathcal{N} = \{1, \ldots, N\}$ of markets (European countries), indexed using the subscript j.[23] The patent on each molecule i has an expiration date, T_i periods into the future, at which point generic alternatives are allowed to enter and profits are driven to zero.[24] The firm's objective is to maximize profits over the life-cycle of their products. We denote the last period of the firm as $T_l = \max_{i \in \mathcal{I}_l} T_i$.

In each period, the firm is solving a two-part problem:

1. In what countries should the products be launched?
2. What prices should be set in each country?

We are interested in understanding strategic launch delays, which are the outcome of the first part of the problem. Of course, the optimal launch strategy will depend on the equilibrium prices that are set in each country. Firms, however, have limited agency in determining these prices, because drug spending is subject to strict government regulation. Therefore, we do not explicitly model the price-setting stage, but instead use a flexible parametric function to predict equilibrium prices.[25]

We start by introducing some notation. Denote the *launch sequence* of firm l as $S_l = \{s_{ij}\}_{j \in \mathcal{N}_i, i \in \mathcal{I}_l}$, where s_{ij} denotes the period of entry of product i in country j. Furthermore, we denote the

[23] Even though in theory all drugs can easily be approved for all countries in the EEA, we allow for the possibility that the drug might not be able to enter in all countries. In some occasions, governments can ban drugs if they are concerned about side effects. Hence, we only assume that a drug can enter in a country if we observe sales in the data.

[24] This assumption can be relaxed in many ways without significantly altering the model. The key result that has to hold is that there are no more strategic incentives to delay once the product has lost patent protection. This is almost certainly the case because once a product loses patent protection, governments can rely on much more effective price-cutting measures, and no longer need to resort to external reference pricing.

[25] We remain agnostic with respect to possible micro-foundations of this function, which do not matter for the results of this paper or for the counterfactual. We show in Appendix A.3 that the equation can be derived from a static Nash Bargaining model.

launch sequence at the end of a given period t as $S_{lt} = \{s_{ijt}\}_{j \in \mathcal{N}_i, i \in \mathcal{I}_l}$ where

$$s_{ijt} = \begin{cases} s_{ij} & \text{if } s_{ij} \leq t \\ 0 & \text{otherwise} \end{cases}$$

Once a product has entered, we assume that it cannot be voluntarily withdrawn.[26] Under this assumption, knowing S_{lt} is enough to know $S_{l\tau}$ for all $\tau < t$. We similarly denote the launch sequence of other firms as S_{-l}. Occasionally, we will also use the shorthand S or S_t to indicate the launch sequences of all firms. Which sequence maximizes profits depends on demand and prices. The demand system and the price-setting equation are primitives that the firm takes as given when making entry decisions. We describe each in turn before specifying the dynamic entry model.

4.1 Demand System

We base demand on the logit random utility model.[27] Markets are defined by country, year, and therapeutic class.[28] We aggregate products within a therapeutic class at the molecule-brand status level. We define three possible brand statuses: originator product (i.e. the brand sold by the patent-holder or main manufacturer), non-originator brand (usually a parallel traded product), and generic. The utility of consumer ℓ, in country j, from consuming drug i (molecule m), belonging to therapeutic class κ, in year t is given by

$$u_{i(m,\kappa)\ell(j)t} = \delta_{ijt} + \nu_{i\ell t} \tag{4}$$

To obtain more realistic substitution patterns we also add a nesting structure at the molecule level. The error term $\nu_{i\ell t}$ is parametrized as

$$\nu_{i\ell t} = \left(\zeta_{m,\kappa} + (1 - \sigma_\kappa)\, \varepsilon_{i\ell t}\right)$$

where m indicates the molecule of drug i, σ_κ lies on the unit interval, $\varepsilon_{i\ell t}$ is distributed according to a standard Extreme Value Type 1 distribution and $\zeta_{m,\kappa}$ is an error term whose distribution satisfies the property that $\nu_{i\ell t}$ is distributed according to an Extreme Value Type 1 distribution as long as $\varepsilon_{i\ell t}$ is also EV1 (Cardell, 1997).[29]

[26] Empirically, we only observe products being withdrawn because the EMA has decided to revoke the marketing authorization upon reviewing post-clinical evidence, or after demand falls for several periods, suggesting that the product is no longer economically or therapeutically viable. In both cases we assume that the choice is not taken by the firm.

[27] Variations on the logit model are commonly used to describe the pharmaceutical market. Duso et al. (2014) and Stern (1996) use a two-level nested logit model to model demand for oral anti-diabetics and a set of four therapeutic classes (gout, sedatives, minor tranquilizers, and oral anti-diabetics) respectively. Dunn (2012) uses a random-effect logit model (micro-BLP) to describe the market for anti-cholesterol drugs using individual data. In our case, we found that adding a more sophisticated nesting structure did not substantially improve model fit (we experimented with adding an upper level at the ATC4 level or at the market level, separating the outside option from all other products). We did not have enough data to implement a random-effect logit model.

[28] For details on the definition and construction of therapeutic classes, see the Online Appendix.

[29] For the nested logit model to make sense, σ_κ must lie on the unit interval. We do not implement this restriction in the estimation, but instead opt to abandon the nesting structure in favor of a simple logit whenever the parameter falls outside the limits set by the theory.

We parametrize δ_{ijt} as

$$\delta_{ijt} = \alpha_{ij} + \beta_i \text{age}_{it} + \eta_i NF_{ijt} + \xi_{ijt} \tag{5}$$

Our specification incorporates two important empirical features of drug demand in our data: heterogeneity in preferences across countries, and growing demand over time.[30] α_{ij} captures a country-specific preference for each drug, which could reflect differences in prescribing guidelines or disease burden.[31] The β_i coefficient allows for the possibility that patients and physicians might learn about new drugs over time. We measure age starting with the approval date of the drug. For non-originator products we also keep track of the number of selling firms as a separate control variable NF_{ijt}.[32] Finally, we add a drug-country-year random shock, ξ_{ijt}, so that the model can fit the data. We do not include a coefficient for price, since we do not observe the price that patients pay. Instead, we include realized demand as a control in the price function, implicitly assuming that any relationship between price and volume sold is mediated by the government. In general, patients in European countries only pay a fraction of the cost of prescription drugs, so any degree of price elasticity that is picked up in the data is likely driven by the government.[33]

Inverting market shares (and normalizing the utility of the outside option to 0) yields the standard estimating equation for nested logit models:

$$\ln\left(\frac{s_{ijt}}{s_{0jt}}\right) = \alpha_{ij} + \beta_i \text{age}_{it} + \eta_i NF_{ijt} + \sigma_m \ln\left(\frac{s_{ijt}}{s_{mjt}}\right) + \xi_{ijt} \tag{6}$$

where s_{0jt} is the share of the outside good, and s_{ijt} and s_{mjt} are the market share of the product, and the overall market share of the molecule nest respectively.[34] We denote the demand function generated by this model as $D_{ijt}\left(S_t, \xi_{jt}^{\kappa}\right)$, where $\xi_{jt}^{\kappa} = \left\{\xi_{ijt}\right\}_{i \in \kappa}$ is the vector of shocks for all products in therapeutic class κ.

[30] Demand for drugs can differ substantially across countries because of heterogeneity in prescribing guidelines, incidence of disease, and patient preferences. Moreover, possibly because drugs are generally considered experience goods (Crawford and Shum, 2005), demand for most products increases over the life-cycle.

[31] Even though the impact of disease burden is mostly reflected through market size, therapeutic classes can sometimes encompass large clusters of disease. For example, oncologics are divided in three large therapeutic classes (targeted therapies, cytotoxics, and hormonal therapies).

[32] Virtually all originator products are sold by a single firm in each country, though the firm is not necessarily the same across countries. However, most molecules face multiple brand and generic competitors, which we aggregate in order to avoid excessive entry and exit.

[33] All European governments provide universal health insurance coverage. Cost-sharing for drugs tends to be very low. Drugs administered in an inpatient setting are usually completely free. Cost-sharing of outpatient prescription drugs is disciplined by a variety of regulations that weaken the relationship between the price paid by the government and that paid by the patient. In a few countries outpatient drugs are also completely free (Netherlands, Scotland). Some countries use copays that are common to all drugs, regardless of price (Austria, Germany, Ireland, Italy, UK). Others have coinsurances but relatively low caps on the amount that each patient can spend each year/month (Belgium, Finland, Spain, Sweden). Finally, countries that have coinsurances without caps either have low coinsurances for the most valuable products (France, Greece, Poland), or allow for a variety of exemptions to protect sick and low-income individuals (Denmark, Portugal) (Barnieh et al., 2014; Panteli et al., 2016; Thomson and Mossialos, 2010).

[34] See Appendix A.1 for the derivation. We think of the outside good as an aggregate of non-drug therapies (doctor visits, surgery, etc.) or drugs in other classes.

4.2 *Price-Setting Equation*

Drug prices are set in negotiations between firms and governments. The exact form of these negotiations is hard to capture in an explicit model. The government is trying to reconcile several goals, such as providing access to valuable medications and rewarding costly innovation, while at the same time facing a budget constraint. Since we do not have any information on the government's objective function we opt for a more agnostic approach, and model prices using a flexible control function.

Our price-setting equation includes two components. The first component is what we call *government price*, p_{ijt}^{gov}. This is the price that is agreed upon between the firm and the government in the absence of reference pricing. We write the government price as a function of product and country fixed effects, as well as three additional control variables that are meant to capture the potential effect of other price-control policies implemented by the government. First, we include an indicator for whether the firm has headquarters in country j. Kyle (2006) shows that this variable is important to determine probability of launch; we include it to check whether we can detect a significant effect on price. Second, we include a flexible function of the number of other molecules available in the same market. There are several reasons why this variable should matter, all of which suggest it should have negative sign: the availability of alternatives should decrease the additional welfare generated by a drug, competitive pressure could bring prices down, and finally, governments sometimes benchmark prices to the lowest price available within a group of substitutable drugs (a practice called internal reference pricing). Third, we include total realized demand for the drug. This variable could also capture a variety of channels, most of which would suggest a negative sign. For example, governments might use soft nudges to steer patients away from expensive drugs to save money. Governments also make widespread use of price-volume agreements meant to prevent budget overshooting. These could take the form of lower prices for drugs whose demand is expected to be higher, as well as volume rebates for drugs with unexpectedly high demand (Carone et al., 2012).

The specification of the government price is

$$p_{ijt}^{\text{gov}}\left(D_{ijt}\left(S_t, \xi_{jt}^{\kappa}\right)\right) = \theta_i \cdot \gamma_j \cdot \exp\left(\beta_Z Z_{ijt} + \beta_D \ln\left(D_{ijt}\left(S_t, \xi_{jt}^{\kappa}\right)\right)\right) \tag{7}$$

where θ_i and γ_j are the product and country fixed effects, Z_{ijt} is the matrix of controls, and $D_{ijt}\left(S_t, \xi_{jt}^{\kappa}\right)$ the realized demand, which depends on the random shocks of the products in class κ: $\xi_{jt}^{\kappa} = \left\{\xi_{ijt}\right\}_{i\in\kappa}$. We interpret this equation as a price-schedule, rather than a set list price.

The second component of the price-setting equation is the *reference price*. The reference price is not directly observed, but reference price functions F_{jt}^{ref} and baskets R_{jt} are reported by various sources.[35] Nonetheless, some details regarding the implementation of these functions require additional assumptions: we need to establish how soon governments see prices that have been set in other countries, and whether ERP is applied before or after volume discounts.

[35] Figure 1 shows the reference baskets and prices for 2012. See the Online Appendix for more details on reference functions in previous years, and on the exact sources of reference price functions and baskets.

We assume that governments see prices with a 1-period lag, and that ERP is applied before volume discounts.[36] Therefore, the reference pricing function that we implement empirically is given by

$$p_{ijt}^{\text{ref}}\left(S_t, D_{ijt}\left(\cdot\right)\right) = F_{jt}^{\text{ref}}\left(\left\{p_{ikt-1}\left(S_t, D_{ijt}\left(\cdot\right)\right)\right\}_{k\in\left(R_{jt}\cap E_{it-1}\right)}\right) \tag{8}$$

where E_{it-1} is the set of countries where product i is sold as of time $t-1$ (this set can be obtained from information in S_t).

To combine these two components, we assume that whenever the governments observes a reference price that is inferior to the government price, the equilibrium price is set as a weighted average of the government price and the reference price. We let the weight be country-specific in order to capture eventual heterogeneity in the application of reference pricing guidelines. The overall price-setting equation is then given by

$$p_{ijt}\left(S_t, D_{ijt}\left(\cdot\right)\right) = \begin{cases} p_{ijt}^{\text{gov}}\left(S_t, D_{ijt}\left(\cdot\right)\right) & \text{if } p_{ijt}^{\text{ref}}\left(\cdot\right) \geq p_{ijt}^{\text{gov}}\left(\cdot\right) \\ \left(1-\mu_j\right) p_{ijt}^{\text{gov}}\left(S_t, D_{ijt}\left(\cdot\right)\right) + \mu_j p_{ijt}^{\text{ref}}\left(S_t, D_{ijt}\left(\cdot\right)\right) & \text{if } p_{ijt}^{\text{ref}}\left(\cdot\right) < p_{ijt}^{\text{gov}}\left(\cdot\right) \end{cases} \tag{9}$$

4.3 Entry Dynamics

We now turn to dynamic choices. The firm operates in a single-agent, discrete-time, finite-horizon environment. Its goal is to maximize profits by choosing the order and timing of entry in each country, conditional on demand and price primitives (over which it has perfect information).[37]

Firms face stochastic shocks in the form of random entry delays. Formally, a delay shock is a binary Bernoulli random variable ρ_{ijt}, with country-specific parameter ψ_j, independently distributed across countries, years, and drugs. If $\rho_{ijt} = 1$, then drug i cannot enter in country j until period $t+1$ (when a new shock is drawn). These shocks help capture variation in delays that cannot be explained through the reference pricing channel. Delays caused by ERP will arise when firms voluntarily decide to withhold one of their products because they expect that doing so will result in higher overall revenues. Launch delays that occur for any other reason will be soaked up by delay shocks.[38]

Each period unfolds as follows. At the beginning, the firm chooses a set of countries where it will send entry applications. We represent this action as a binary vector $A_{lt} = \left\{a_{ijt}\right\}_{j\in\mathcal{N}_l, i\in\mathcal{I}_l}$ where $a_{ijt} = 1$ whenever the firm chooses to launch product i in country j. More generally, we

[36] We think this is the most natural sequence, since ERP is used to set the initial price, while volume discounts can only be applied at the end of the year.This sequence implies that for the purposes of ERP, governments use the initial prices (i.e. *before* volume adjustments), though in our data we observe the final price (inclusive of eventual volume discounts). As a result, we calculate prices without the volume component when calculating reference prices.

[37] Notice that revenue and profits are equivalent, since the model assumes that all costs are zero.

[38] The main source of delays other than ERP is probably the time required to review price and reimbursement applications and to complete price negotiations. However, other sources may exist as well: firms may need to wait for certain negotiations to be resolved before engaging in additional launching because of capacity constraints to their negotiating workforce; or countries may block the entry of products they consider potentially dangerous for a number of years.

denote *a strategy* for firm l in the extended-form problem as a map

$$\mathcal{A}_{lt} : \mathcal{S} \to \{0,1\}^{T_l - t + 1}$$

where $\mathcal{S}$ is the set of all possible values of the launch sequence S_l, and $\mathcal{A}_{lt}$ generates a set of binary vectors A_{lt} with an action profile for each period until T_l. The launch sequence of the firm as of period t represents the state variable of the problem. While the launch sequence of other firms has the potential to affect expected revenue (by stealing market share and potentially affecting prices), we assume that strategies are not conditional on the actions of other firms.[39] After the vector of binary shocks for the current period $\{\rho_{ijt}\}$ is realized, the value of the state variable updates, governments set prices, and finally, products are sold, and profits are realized.

The firm's problem at time t is to pick a strategy to maximize

$$V_t\left(S_{lt-1}, S_{-lt-1}\right) = \max_{\mathcal{A}_{lt}=\{A_{l\tau}(S_{l\tau-1})\}_{\tau=t}^{T}} \sum_{S_l} \left(\sum_{S_{-l\tau}} \left(\sum_{\tau=t}^{T} \beta^{\tau-t} \Pi_\tau\left(S_l, S_{-l}\right) \right) \cdot \mathcal{P}\left(S_{-l} \middle| S_{-lt-1}, \mathcal{A}_{-lt}\right) \right) \cdot \mathcal{P}\left(S_l \middle| S_{lt-1}, \mathcal{A}_{lt}\right) \tag{10}$$

where β is the discount factor, $\Pi_\tau\left(S_l, S_{-l}\right)$ is the expected period profit of the firm, for a given realization of the launch sequence (both the own sequence and the sequence of competitors), and $\mathcal{P}\left(S_l \middle| S_{lt-1}, \mathcal{A}_{lt}\right)$ and $\mathcal{P}\left(S_{-l} \middle| S_{-lt-1}, \mathcal{A}_{-lt}\right)$ are the probabilities of S_l and S_{-l} conditional on S_l and S_{lt-1}, for given strategies $\mathcal{A}_{lt}$ and $\mathcal{A}_{-lt}$ of the firm and its competitors.

The expected period payoff is defined as

$$\Pi_\tau\left(S_l, S_{-l}\right) = \sum_{i \in \mathcal{I}_l} \mathbb{E}_{\xi^r_{jt}} \left[\sum_{j \in S_{l\tau}} p_{ij\tau}\left(S_t, D_{ijt}\left(\cdot\right)\right) D_{ij\tau}\left(\cdot\right) \right] \tag{11}$$

where the expectation is taken over the possible realizations of the stochastic error in the demand system.

5 Estimates of Demand and Price Primitives

5.1 Demand Estimation

We estimate demand from equation 6. The independent variables can be easily constructed from the IMS data, with the exception of age, for which we use approval date from the EMA or the Heads of Medicines Agencies.[40] To measure market size, we use data from the Global

[39] This assumption is necessary to obtain identification from the model. Allowing firms to react to each other is desirable, but makes counterfactual scenarios hard to compute, since our data does not allow us to make inferences about reactions off the equilibrium paths.

[40] We manage to match over 90% of molecules in our main sample to an approval date. In the handful of cases for which we do not have a match, we use the fifth-earliest launch date from the IMS MIDAS database as a proxy for the approval date. In the sample we were able to match, the fifth-earliest launch occurs after the approval date in 95% of the cases. Notice that any measurement error in the actual approval date will not impact demand estimation, since it will be absorbed by the country-drug fixed effect.

Burden of Disease Study. We use a map from ATC4 to GBD indication constructed by Costinot et al. (2016) to calculate the number of patients that might potentially use drugs in a certain therapeutic class.[41] We then scale up the number of patients to obtain a number for standard units and construct market shares from data on sales volumes.[42]

Identification of demand system parameters

Two potential identification issues arise. First, $\ln\left(\frac{s_{ijt}}{s_{mjt}}\right)$ (i.e. the within-molecule market share of product i) is correlated with the error term ξ_{ijt}, so we need instruments to recover a consistent value for σ_m. We use three. First, we use the total number of other firms that are selling the same product (distinguishing between brand/parallel traded products and generics). This is because in a logit model, the within-molecule share will be mechanically related to the number of alternative options. Second, we use years since the patent on molecule m expired. This instrument is motivated by the fact that market shares tend to shift to generic manufacturers over time after loss of exclusivity. Third, we use the average within-molecule market share of parallel traded products for other molecules within the same country j.[43] This instrument is meant to capture the average propensity of a government to shift individuals towards parallel traded products.

The second issue is that firms might be able to observe ξ_{ijt} prior to entry. This leads to a classic selection problem common to many IO settings: countries where entry is recorded would have unobservably high values of ξ_{ijt}, leading to a biased estimator. In practice however, there are several attenuating circumstances that suggest selection is a second-order concern in this case. First, firms never exit voluntarily, so we do not need to worry about exit selection. Second, since we control for drug-country-specific preferences, our model will pick up the average preference of each country. The remaining concern is then that demand prior to entry could be lower than our model predicts, because firms may wait until demand reaches a certain threshold before entering. If that were the case, we would expect our α_{ij} coefficients to be biased upwards. In the dynamic estimation, a higher α_{ij} coefficient makes firm i more likely to enter in country j. Therefore, this type of selection bias would lead us to underestimate the extent of strategic delays.[44] We conclude that if unobserved selection does exist, it should act in the direction of preventing us from finding a result.

Results

We estimate a separate equation for each of the 109 therapeutic classes that contain at least one of the products in our main sample using linear regressions with instruments. The actual value of the estimated coefficients does not have an intuitive economic interpretation. Instead, we are more concerned with the ability of the model to fit and predict demand accurately. Virtually

[41] We thank the authors of the paper for sharing the map with us ahead of publication.

[42] The scaling number is chosen to be the smallest number such that the outside option has at least 1% market share in all countries and years. In that sense, our estimate can be thought of as a lower bound on the actual market size. We pick a different scaling number for each therapeutic class.

[43] We identify as parallel traders all non-originator firms who sell a product sharing the same molecule and product name as the original product.

[44] If potential demand were higher before entry, that might lead us to overestimate delays due to strategic behavior. However, this type of selection runs counter to the main intuition for the selection mechanism in this case.

all regressions achieve a coefficient of determination of at least 0.8. This is both a result of our very flexible model, and a reflection that demand for most drugs tends to be well-behaved, without many fluctuations.

5.2 *Price and ERP parameters*

We estimate the pricing function using equation 9. To fit the model, we assume that prices are subject to measurement error. This error could come from a variety of sources: our prices are yearly averages, which could shroud higher frequency fluctuations; IQVIA collects their data from pharmacies and hospitals, so its precision depends on reporting systems that may be less than accurate. For simplicity, we assume that η_{ijt} is i.i.d. across countries, drugs and years, and do not include any additional sources of error.[45] Since our price function is multiplicative, we also assume that the measurement error is multiplicative (i.e. additive in logs). Denotingp_{ijt} as the model-predicted price, and p^o_{ijt} as the observed price we obtain

$$\ln\left(p^o_{ijt}\right) = \begin{cases} \ln\left(p^{\text{gov}}_{ijt}\left(\cdot\right)\right) + \eta_{ijt} & \text{if } p^{\text{ref}}_{ijt}\left(\cdot\right) \geq p^{\text{gov}}_{ijt}\left(\cdot\right) \\ \ln\left(\left(1-\mu_j\right) p^{\text{gov}}_{ijt}\left(\cdot\right) + \mu_j p^{\text{ref}}_{ijt}\left(\cdot\right)\right) + \eta_{ijt} & \text{if } p^{\text{ref}}_{ijt}\left(\cdot\right) < p^{\text{gov}}_{ijt}\left(\cdot\right) \end{cases}$$

Our estimation routine searches the vector of parameters that minimizes the difference between the model prediction and the data. Since the price function includes a fixed effect for each product and country (there are 481 distinct products in our main sample), the total number of parameters is quite large, and estimation matching log price levels would be quite slow. To improve the speed and efficiency of the procedure we match log differences in price (the key property of this ratio is that it does not depend on the product fixed effect θ_i, which we prove in Appendix A.2). Since log differences do not depend on the product fixed effect θ_i, this strategy drastically reduces computation time. We match $\ln\left(\frac{p_{ijt}}{p_{ikt+1}}\right)$, where j and k are two (randomly selected) countries, and we look at the difference in the price of product i in these two countries over consecutive periods.[46] The estimating equation in differences is

$$\ln\left(\frac{p^o_{ijt}}{p^o_{ikt+1}}\right) = \ln\left(\frac{p_{ijt}\left(\cdot\right)}{p_{ikt+1}\left(\cdot\right)}\right) + \eta_{ijt} - \eta_{ikt+1}$$

and our routine minimizes the sum of squares of the two error terms:

$$O\left(\gamma_j, \mu_j, \beta_Z, \beta_D\right) = \sum_{i,j,k,t}\left[\ln\left(\frac{p_{ijt}\left(\cdot\right)}{p_{ikt+1}\left(\cdot\right)}\right) - \ln\left(\frac{p^o_{ijt}}{p^o_{ikt+1}}\right)\right]^2 \tag{12}$$

We restrict the μ_j parameters to the unit interval. A negative value of μ_j does not make sense; a value of μ_j greater than 1 — while theoretically possible — raises the possibility of negative prices in counterfactual predictions, which is undesirable. Once we have obtained estimates

[45] Introducing any other source of error in the price would potentially lead to insurmountable estimation issues due to the propagation of this error through the reference pricing channel.

[46] We compare prices of different countries to avoid differencing out the country fixed effect. In order to minimize the number of observations lost through the differencing process, we occasionally consider $\frac{p_{ij2012}}{p_{ik2002}}$ as a moment. By doing so we only lose one observation per drug, instead of one observation per drug-country.

for $(\hat{\gamma}_j, \hat{\mu}_j, \hat{\beta}_Z, \hat{\beta}_D)$, we plug them into the price equation to calculate $\hat{\theta}_i$.

Identification of the pricing equation

The main threat to identification in the estimation of the pricing equation is the potential presence of cross-country correlations in the error term η_{ijt}. For example, in the case of an EU-wide demand or cost shock we would expect prices of all drugs to be affected similarly. The fear is that our model might erroneously interpret this as a result of reference pricing. To make sure this does not happen, instead of calculating reference prices using *observed* prices in the previous period, we we calculate them using *predicted* government prices. As a result, μ is identified through comovements in observed prices and predicted reference prices. This ensures that η_{ijt} does not affect reference prices, and therefore cannot generate a spurious correlation between prices of different countries.

The variation in predicted reference prices that identifies μ_j comes from two possible sources. First, the reference price will change if the value of Z_{ijt} changes in any of the countries where the product is already available. Second, it will change when the drug launches in a new country. Hence, one of the key sources of variation that pins down μ_j is the exact timing of entry of drugs across countries. It is important that this variation is plausibly exogenous with respect to changes in prices in country j. What helps us is the fact that the entry application process injects significant randomness in the timing of entry of new products. The delays generated by government negotiations can be significant and are likely orthogonal to prices and strategic considerations. Indeed, while most drugs follow relatively similar launch patterns, there is substantial variation in the exact order of drug launches, to the point where virtually no two drugs follow the same exact entry sequence. The fact that entry sequences have a stochastic component also helps us identify the other components of the pricing equation. For virtually every country there are at least a few drugs that begin their launch sequence there. Since governments cannot observe a reference price at the beginning of the launch sequence, this variation helps us identify the components of the government price function.

In calculating the change in the reference price we also leverage the fact that we know the reference function and basket, and the assumption that governments see prices with a one-period lag. These two factors affect when the reference price reacts to a change in prices abroad and the degree to which it is affected. For example, a drug's Italian reference price will adjust in the period after a drug is launched in Poland, while the French reference price will not — Poland is in Italy's reference basket, but not in France's. Moreover, the extent to which the Italian reference price moves will depend on the number of countries where the product is alreay available: Italy references 24 countries and uses average as their reference function. Finally, since Italy is one of the countries referenced by France, the French reference price may adjust two periods later to reflect any eventual movements in the Italian price caused by the launch in Poland.

Results

We report price estimation results for the vector of parameters $(\hat{\gamma}_j, \hat{\mu}_j, \hat{\beta}_Z, \hat{\beta}_D)$. Since our estimation is in logs, we report coefficients for $\ln(\gamma_j)$, which are more easily interpretable as

proportional changes in relative terms with respect to a benchmark (in this case, the omitted coefficient used as a benchmark is $\ln(\gamma_{\text{GERMANY}})$).

The first column of Table 3 shows the coefficients for $\ln(\gamma_j)$. The point estimates roughly match our intuition: countries with lower income tend to pay lower prices, with a couple of exceptions. Poland and Hungary have higher coefficients than many other countries with higher income. However, both countries use the minimum price in Europe as their reference, and the coefficients on μ_j for both are very close to 1. This suggests that the government may be willing to grant higher prices, knowing that reference rules will bring them down very quickly.[47]

The second column shows estimates for μ_j, which is the coefficient measuring how strictly each country adheres to its own ERP guidelines. We observe significant heterogeneity across countries in this respect. Seven countries have coefficients above 0.9, suggesting that reference pricing guidelines are followed closely. Four other countries (including Italy and Spain) have coefficients above 0.8. The remaining 8 countries who use ERP all have coefficients below one-third, which suggests that they either do not follow their guidelines all that closely, or that they apply them only to selected drugs. In particular Switzerland does not appear to use reference pricing at all, with a coefficient that is almost exactly zero. We are not able to identify a coefficient for Slovenia because the model estimates that its government price is always below the reference price (Slovenia references Austria, France, and Germany, which are all countries with much higher price levels).

Almost all the coefficients on the control variables in the welfare function behave as expected. Higher quantity sold is associated with lower prices. Prices tend to be approximately 4% higher in countries where the firm has headquarters. Finally, having a higher number of competitors in the same class is associated with lower prices, though the relationship appears to be nonlinear and noisy.

One important result that emerges from the analysis is that the price level (as indicated by $\ln(\gamma_j)$) in Western European countries that follow reference pricing closely is only marginally higher than the price level of Eastern European countries. This suggests that in the current equilibrium firms may be under pressure to keep prices higher in Eastern European countries to reduce the size of the externality generated by ERP.

5.3 Simulation-Based Evidence of Optimal Delays

Our price estimation results suggest that ERP does indeed affect equilibrium prices, but are the externalities strong enough to generate delays? We test this by simulating firm revenue from two groups of strategies and compare it to the payoff obtained when launching everywhere right away.

The first group of strategies, denoted as $\mathcal{A}^{kr}$, consists of withholding a single product from

[47] Luxembourg and Norway (the two richest countries in Europe) are also outliers. Norway is relatively isolated in the reference pricing matrix, because it is only referenced by Finland. Hence granting a lower price to Norway might carry relatively little consequences for firms. This effect is not captured explicitly by our model, but is incorporated in the country fixed effect. Luxembourg is harder to explain, though its status as a relatively small country might give rise to all sort of irregularities and exceptions. Many drugs do not even record sales in Luxembourg, so it is possible that selection might play a role here.

Table 3: PRICE ESTIMATION RESULTS

Country	$\ln(\gamma_j)$		μ_j	
Austria	-0.099	[-0.134,-0.071]	0.199	[0.022,0.444]
Belgium	-0.123	[-0.156,-0.093]	0.119	[0.010,0.446]
Bulgaria	-0.230	[-0.317,-0.145]	0.987	[0.432,0.998]
Denmark[b]	-0.068	[-0.102,-0.046]	0	
Estonia	-0.170	[-0.380,-0.128]	0.994	[0.249,0.998]
Finland	-0.126	[-0.165,-0.098]	0.332	[0.018,735]
France	-0.098	[-0.130,-0.061]	0.098	[0.008,841]
Germany[a,b]	0		0	
Greece	-0.186	[-0.216,-0.035]	0.987	[0.841,0.999]
Hungary	-0.145	[-0.240,-0.082]	0.991	[0.326,0.998]
Ireland	-0.089	[-0.127,-0.062]	0.229	[0.018,0.840]
Italy	-0.191	[-0.221,-0.110]	0.854	[0.219,0.994]
Latvia	-0.201	[-0.317,-0.151]	0.805	[0.013,0.941]
Lithuania	-0.204	[-0.359,-0.151]	0.996	[0.014,0.999]
Luxembourg[c]	-0.200	[-0.247,-0.176]	0	
Netherlands	-0.183	[-0.220,-0.151]	0.293	[0.073,0.826]
Norway	-0.143	[-0.186,-0.111]	0.846	[0.107,0.981]
Poland	-0.119	[-0.238,0.025]	0.904	[0.239,0.989]
Portugal	-0.176	[-0.220,-0.138]	0.093	[0.012,0.842]
Romania	-0.253	[-0.351,-0.105]	0.994	[0.204,0.999]
Slovenia[d]	-0.232	[-0.272,0.204]	/	
Spain	-0.159	[-0.186,-0.131]	0.874	[0.151,0.975]
Sweden[b]	-0.102	[-0.131,-0.080]	0	
Switzerland	0.007	[-0.024,0.032]	0.020	[0.003,0.080]
UK[b]	-0.201	[-0.233,-0.179]	0	
Controls				
Log quantity sold	-0.019	[-0.024,-0.017]		
Home-Firm Indicator	0.041	[0.011,0.079]		
At least 1 other molecule in class	-0.043	[-0.079,0.028]		
At least 2 other molecules in class	0.013	[-0.024,0.045]		
At least 5 other molecules in class	-0.029	[-0.063,0.007]		
At least 10 other molecules in class	-0.003	[-0.031,0.016]		

[a] The price level is normalized to Germany's.

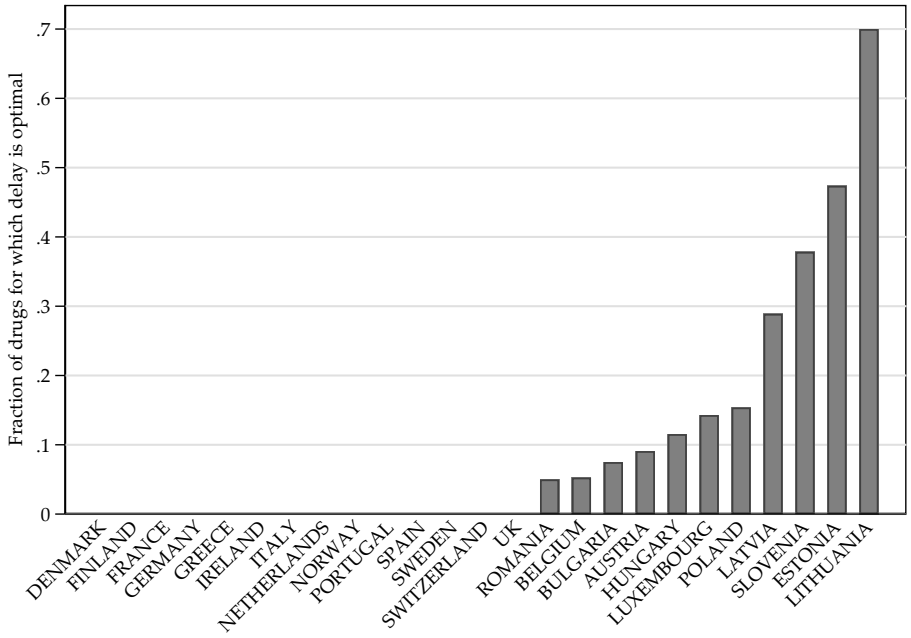
[b] Denmark, Germany, Sweden, and the UK do not use reference pricing during the period 2002-2012, so we set μ_j to zero.

[c] Luxembourg references the price of the country of origin of the drug. We don't know country of origin, so we simply assume that μ_j is equal to 0.

[d] On occasion, a country's price level is such that its reference price never binds. In these cases the coefficient is undetermined.

This table reports coefficients from the price estimating equation (equation 12). $\ln\left(\gamma_j\right)$ is the country fixed effect, in log terms, normalized with respect to the coefficient of Germany. μ_j represents the weight assigned to the reference price. Values close to 1 mean that reference price is followed closely, while values close to 0 mean that reference price does not matter (μ_j is restricted to the unit interval by design). See Section 4.2 for a detailed description of the price function and control variables. Standard errors are calculated using nonparametric bootstrap with sampling at the drug level.

Figure 6: OPTIMALITY OF STRATEGIC DELAYS BY COUNTRY



This figure plots the fraction of drugs for which delaying in each country is optimal. We use a sample of 87 drugs whose patent expires on or before December 31st, 2012. By the time our sample started, in 2002, launches in many countries had already occurred for some of these drugs. Hence, for each country j we estimated the y-axis variable using the subsample of drugs that had not already entered in country j before 2002. This implies that the sample used is slightly different for each country.

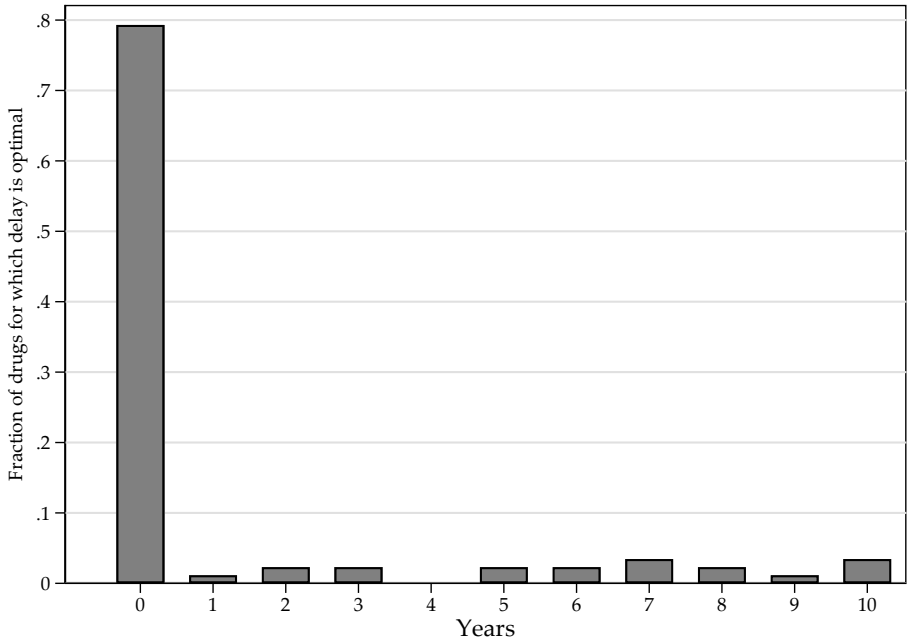
country k for r periods. We simulate the expected revenue of these strategies for all the 87 drugs in our dynamic estimation sample and calculate the fraction of drugs for which delaying in country j is optimal for at least one value of r.[48] Figure 6 plots the results. We find that these types of delaying strategies are optimal for all Eastern European countries in at least some cases. Only three countries outside of Eastern Europe would experience delays according to this simulation: Austria, Belgium, and Luxembourg. Luxembourg does not have particularly low prices, but it's a very small market, so it's not surprising that for some drugs it would be optimal to exclude it. The model also predicts a delay for one drug in Austria, and one in Belgium. In both cases, the drug in question was indeed launched with a long delay in both countries and earned a very small amount in sales. Even though both Belgium and Austria tend to have high price levels relative to most other countries, they can affect the prices of countries with higher price levels (for example Ireland).[49]

Figure 6 is remarkably consistent with the reduced-form patterns described in Section 2.3 (Figure 2). No information about delays or launch sequences was used in generating the figure,

[48] In some cases, drug i is already in country j in 2002, which makes it impossible to simulate revenue under a counterfactual delay. We exclude those drugs from our calculations.

[49] Delaying entry in Romania and Bulgaria is optimal for a surprisingly low number of drugs, given our price and demand estimates. This is because prior to acquiring EU membership in 2007, Bulgaria and Romania where only referenced by a few other Eastern European countries. Thus, our model predicts that entering in Romania and Bulgaria only has a small effect on prices prior to 2007. In the data, we do not find a significant difference in the average delays before and after EU entry in these two countries. This is not too surprising however, as entry in the EU also removes several bureaucratic obstacles that might have generated idiosyncratic delays. Hence, the net effect of EU membership on launch delays could be close to 0.

Figure 7: STRATEGIC DELAYS IN EASTERN EUROPEAN COUNTRIES



In this figure we consider a set of strategies where entry in Eastern European countries is delayed for t periods, where t is between 0 and 10. We count the number of drugs for which each strategy is optimal and plot it in a histogram (the y-axis displays the fraction of drug, rather than the level).

which is purely based on price and demand estimates. This confirms the intuition that firm behavior adheres to the incentives laid out by the regulatory environment and supports the idea that our model captures the relevant features of this market.

In the next set of strategies, we focus on Eastern Europe as a block, and simulate revenue from strategies that delay entry in all Eastern European countries.[50] We plot the distribution of optimal delays for our 87 drugs in Figure 7. We find that for around 20% of drugs, delaying in all Eastern European countries as a block is preferable.

These simulations suggest that firms are better off when their products are not available in Eastern European countries right away. At the same time, we find almost no evidence that firms have any strategic incentive to delay entry in countries in Western Europe (this includes even countries with relatively lower income levels like Greece and Portugal).

6 DYNAMIC ANALYSIS

With estimates for demand and price primitives in hand, we now turn our attention to the parameters governing idiosyncratic delays. These parameters, which may differ across countries, capture the probability that an application may be delayed.

Traditional estimation methods for dynamic entry models are unfeasible in our case for three reasons. First, the model does not have an analytic solution due to the complicated net of price externalities. Second, the state space (N^T possible entry sequences for N countries and T

[50] As with the previous simulation exercise, sometimes drug are already available in some Eastern European countries. We delay entry only in countries where the product has not already entered.

periods) is too large to solve the problem numerically. Third, firm strategies are unobserved: we see when drugs enter, but not when firms send applications. In other words, in the observed launch sequences, an application that was not sent is observationally equivalent to an application that was rejected. To overcome these obstacles, we approach the problem using a partial identification approach.

We use two sets of moment inequalities to construct bounds on the probability of idiosyncratic delays. The first set of moment inequalities uses data on entry and approval dates and exploits the intuition that firms can only apply after they have received marketing approval for a product. Hence, the average probability of an idiosyncratic delay is bounded above by the overall probability of delay implied by the data. The second set of inequalities uses information contained in the expected revenue of the observed entry sequence. The estimator rejects parameter values for which we can find strategies that yield higher expected revenue than what firms obtained in the data.

In constructing the second set of moment inequalities we introduce a methodological contribution by developing a new method to build moment inequalities that does not require observing the firm's strategy. We show that, conditional on some assumptions, simply observing the total revenue earned by the firm is sufficient to derive restrictions on model parameters. Our inequalities compare observed revenue to simulated revenue using arbitrary strategies and reject parameter values for which the simulated revenue is higher.

6.1 Moment Inequalities based on Entry Data

We use entry and approval data to recover an upper bound on the delay parameters. Our distributional assumption on the random delay shocks is that in each period the probability of an application for entry in country j being delayed is ψ_j. Let $\left(1-\bar{\psi}_j\right)$ be the probability that product i will enter in country j in a given year. This is the combination of the probability that the firm will apply, times the probability of the application being accepted. Hence, for all j, $\bar{\psi}_j \geq \psi_j$.

To estimate $\bar{\psi}_t$ we can simply calculate the probability of a delay by using data on approval dates and launch dates: suppose that a product approved in year 0 enters France in year 2. Then we are going to assign an expected probability of entry of $1/3$: the product had three opportunities to enter — in years 0, 1, and 2 — and registered one success, in year 2.

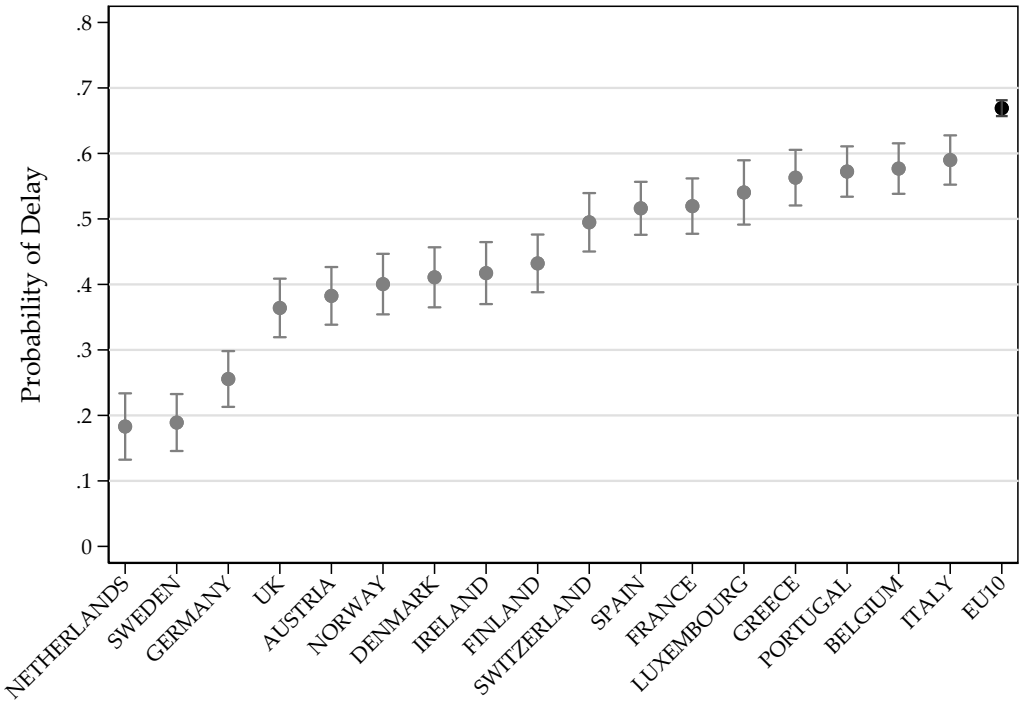
Figure 8 plots the coefficients $\bar{\psi}_j$ for all European countries, together with confidence intervals.[51] We estimate a single coefficient for Eastern Europe — denoted as ψ_{EU10} — for consistency with the moment inequality estimation (see the next section).[52] The range of these parameters varies between 0.2 and 0.6 for Western European countries, while the coefficient for Eastern Europe is 0.669.[53]

[51] We follow instructions laid out by Brown et al. (2001) in building approximate confidence intervals for these parameters.

[52] EU10 is the shorthand for Eastern European countries that are part of the European Union. In our data we have 8 Eastern European countries (instead of 10), because we are missing sales data for Czech Republic and Slovakia.

[53] Under the assumptions of the model, one could also obtain a tighter bound by choosing a partition P of the set of all drugs, calculating a subset-specific upper bound $\bar{\psi}_j^p$ for all $p \in P$, and choosing $\bar{\psi}_j^{p,\min} = \min_p \bar{\psi}_j^p$ to be the upper bound. We show in the Online Appendix that doing so can yield a lower upper bound for $\bar{\psi}_j$. However, we only include the estimate from the full sample in our main results, as this exercise relies heavily on the assumption

Figure 8: PROBABILITY OF OVERALL DELAYS BY COUNTRY



In this figure we plot the overall probability of a delay in each country in Europe. We aggregate Eastern European countries as a whole for consistency with the moment inequality estimation. The plot includes 95% confidence interval bands for each parameter.

6.2 Moment Inequalities based on Revenue Data

Identification in moment inequality is generally based on the idea that the firm's strategy contains some information about the true value of the unknown parameter of a model. In particular, the strategy of the firm will only be optimal for a certain range of values of the parameter. For values outside that range we should be able to find strategies that would earn higher revenue, which in turn allows us to rule those values out.

In our setting firm strategies are unobserved (e.g. if we observe entry in France with a 2-year delay, we don't know whether the firm applied starting in year 0, year 1, or year 2). As a result, we are unable to compute the expected revenue of the firm's observed strategy for arbitrary values of the unknown parameter. However, the data tells us what the firms did earn. Under the assumption that the model is well-specified, this corresponds to the revenue the firm earns at the true value of the parameter. We use this information to obtain identification and build inequalities in a way that is similar to the rest of the literature: we compare the revenue earned by the firm in the observed data and calculate counterfactual revenue under alternative strategies for arbitrary values of the parameter.

The loss of information that comes from not observing the optimal strategy means that our inequalities will not necessarily hold for individual firms: the revenue we observe depends not only on strategies, but also on the realization of the delay shocks. If a firm incurs in longer-

that ψ_j is constant across drugs.

than-expected delays, this will decrease their revenue and we may end up finding a violation of the inequality even at the true parameter value. To address this issue, we aggregate across drugs and instead match the average expected revenue across all firms. Doing so will remove the noise generated by the realization of the delay shocks.

Formally, let $\mathcal{A}^{\star}_{lt}$ denote the optimal strategy of firm l starting in period t. By definition, $\mathcal{A}^{\star}_{lt}$ is the solution to the dynamic programming problem of the agent as expressed in equation 10. In other words, for all possible strategies $\mathcal{A}'_{lt}$, $\mathcal{A}^{\star}_{lt}$ satisfies

$$\tilde{V}_t\left(\mathcal{A}^{\star}_{lt}; S_{lt-1}, S_{-lt-1}\right) \geq \tilde{V}_t\left(\mathcal{A}'_{it}; S_{lt-1}, S_{-lt-1}\right) \tag{13}$$

where

$$\tilde{V}_t\left(\mathcal{A}_{lt}; S_{lt-1}, S_{-lt-1}\right) = \mathbb{E}\left[\sum_{\tau=t}^{T} \beta^{\tau-t} \Pi_{\tau}\left(S_l, S_{-l}\right) \middle| \mathcal{A}_{lt}, S_{lt-1}, S_{-lt-1}\right] \tag{14}$$

is the expected payoff of the agent conditional on playing strategy profile $\mathcal{A}_{lt}$ (the expectation is taken over the possible realizations of the launch sequences of all products, as in equation 10).

Under the assumption that firms maximize expected returns we can use equation 13 to build "revealed preference" inequalities as in Pakes et al. (2015). In the data we observe a certain number of firms with molecules in specific therapeutic classes (each firm-class combination constitutes one observation for us, since molecules in different therapeutic classes do not affect each other). The *expected* payoff of the observed launch sequence S^o_l of firm l starting at period t is

$$V_t\left(S^o_l, S^o_{-l}\right) = \sum_{\tau=t}^{T} \beta^{\tau-t} \Pi_{\tau}\left(S^o_l, S^o_{-l}\right)$$

Notice that $\Pi_{\tau}\left(S^o_l, S^o_{-l}\right)$ is an object that we can recover using simulation, since it only depends on the shock $\xi^{\kappa}_t = \left\{\xi^{\kappa}_{jt}\right\}$. This error is unobserved by both the firm and the econometrician by assumption.[54] We recover the distribution of ξ^{κ}_t from demand estimation and use it to simulate ξ^{κ}_t. Then, the average of $V_t\left(S^o_l, S^o_{-l}\right)$ converges to the average of the expected payoff of each firm when playing the optimal strategy.

Theorem 1. *For any $\varepsilon > 0$, we can find M' such that*

$$\frac{1}{M}\sum_{i=1}^{M}\left(V_t\left(S^o_l, S^o_{-l}\right) - \tilde{V}_t\left(\mathcal{A}^{\star}_{lt}; S_{lt-1}, S_{-lt-1}\right)\right) < \varepsilon$$

for all $M > M'$.

We provide a rigorous proof of this Theorem in Appendix A.4 and only discuss the intuition behind the result here. Notice that if firms are playing the optimal strategy, then the observed launch sequences are draws from the probability distributions $\mathcal{P}\left(S_l \middle| S_{lt-1}, \mathcal{A}^{\star}_{lt}\right)$ and $\mathcal{P}\left(S_{-l} \middle| S_{-lt-1}, \mathcal{A}^{\star}_{-lt}\right)$. This distribution need not be the same for all firms: each may face a different initial state S_{lt-1}, and different demand or price primitives. However, since the random shocks used in the model are independently distributed across drugs, we can apply a gener-

[54] There is no additional error from price since the residual is assumed to be measurement error.

alized version of the law of large number for non-identical independent random variables to argue that the sample average across firms converges to the average expected payoff.

Theorem 1 suggests that we can write moment inequalities based on the average payoff obtained by the firms in our sample:

$$\frac{1}{M}\sum_{i=1}^{M}\tilde{V}_t\left(\mathcal{A}_{lt}^{\star};S_{lt-1},S_{-lt-1}\right)\geq\frac{1}{M}\sum_{i=1}^{M}\tilde{V}_t\left(\mathcal{A}_{lt}';S_{lt-1},S_{-lt-1}\right) \tag{15}$$

The left-hand side of this inequality is approximated by $\frac{1}{M}\sum_{i=1}^{M}V_t\left(S_l^o,S_{-l}^o\right)$. To compute the right-hand side we use simulation. The expected payoff of any deviation $\mathcal{A}_{lt}'$ can be written as an integral over the possible realization of the launch sequence S_l, holding S_{-l} fixed [55]

$$\tilde{V}_t\left(\mathcal{A}_{lt}';S_{lt-1},S_{-lt-1}\right)=\sum_{S_{-l}}\sum_{S_l}\sum_{\tau=t}^{T}\beta^{\tau-t}\Pi\left(S_l,S_{-l}\right)\mathcal{P}\left(S_l|\,S_{lt-1},\mathcal{A}_{lt}'\right)\mathcal{P}\left(S_{-l}|\,S_{-lt-1},\mathcal{A}_{-lt}^{\star}\right)$$

We approximate $\tilde{V}_t\left(\mathcal{A}_{lt}';S_{lt-1},S_{-lt-1}\right)$ by simulating the distribution of $\mathcal{P}\left(S_l|\,S_{lt-1},\mathcal{A}_{lt}'\right)$. For a given guess of the parameter vector ψ_{EU10}, and for each drug $i\in\mathcal{I}_l$ draw $\left\{\nu_{ijt}^r\right\}_{r=1}^{nsim}$ and use them to calculate simulated entry paths $\left\{S_l^r\right\}$. The average simulated payoff is

$$V_t^{sim}\left(\mathcal{A}_{lt}';S_{lt-1},S_{-lt-1},\psi_{EU10}\right)=\frac{1}{nsim}\sum_{r=1}^{nsim}\left[\sum_{\tau=t}^{T}\beta^{\tau-t}\Pi\left(S_l^r,S_{-l}\right)\right]$$

The difference between $\tilde{V}_t\left(\mathcal{A}_{lt}';S_{lt-1},S_{-lt-1}\right)$ and $V_t^{sim}\left(\mathcal{A}_{lt}';S_{lt-1},S_{-lt-1},\psi_{EU10}^0\right)$ — for the true parameter vector ψ_{EU10}^0 — is simulation error, which can be eliminated by choosing *nsim* large enough, and the error in the realization of S_{-lt-1} (which is only one of many possible draws). When we aggregate across firms, the error in S_{-lt-1} will disappear for a large enough sample of drugs.

We define empirical moment conditions as

$$\mu\left(\mathcal{A}_{lt}',\psi_{EU10}\right)=\min\left\{0,\frac{1}{M}\left(\sum_{i=1}^{M}V_t\left(S_l^o,S_{-l}^o\right)-V_t^{sim}\left(\mathcal{A}_{lt}';S_{lt-1},S_{-lt-1},\psi_{EU10}\right)\right)\right\}$$

The estimation identifies a set Ψ^I of parameters that satisfy

$$\Psi^I=\left\{\psi_{EU10}\,:\,\sum_{\mathcal{A}_t'}\mu\left(\mathcal{A}_{lt}',\psi_{EU10}\right)=0\right\}$$

Results Based on our simulation results, we assume that firms never optimally delay entry in Western Europe (i.e. $\bar{\psi}_j=\psi_j$ for all countries outside Eastern Europe). Since moment inequalities work best with a small number of parameters, we assume that the probability of delay is the same in all Eastern European countries. This parameter captures the probability of

[55] We cannot simulate S_{-l} because we do not observe strategies. However, each observation of S_{-l} is drawn from the probability distribution of $\mathcal{P}\left(S_{-l}|\,S_{-lt-1},\mathcal{A}_{-lt}^{\star}\right)$, so when we average across firms, we will approximate the right distribution. See the proof of Theorem 1 in Appendix A.4.

incurring a delay in Eastern European countries due to forces outside of the firm's control. We estimate an identified set for this parameter by building moment conditions based on equation 15, and checking what range of parameters satisfies it for a series of possible strategies $\mathcal{A}'_{lt}$.

Calculating $V_t\left(S^o_l, S^o_{-l}\right)$ and $V^{sim}_t\left(\mathcal{A}'_{lt}; S_{lt-1}, S_{-lt-1}, \psi_{EU10}\right)$ does not require observing drugs from their original launch, but it does require observing them until period T_l. Hence we perform this analysis on the dynamic sample (see Table 1 for details).

We let our simulation analysis (Section 5.3) guide our choice of counterfactual strategies. For each drug i, the simulations can tell us the optimal entry period country j, $s^\star_{ij}$ (conditional on entering in all other countries right away). We then take $s^\star_{ij}$ for all drugs i and countries j and build strategies where firms send entry applications k periods in advance, for $k = 3$. On top of that, we also test a variety of simpler strategies, such as applying right away in all countries, or applying k periods before observed entry s^o_{ij}. These strategies yield looser bounds.

The identified set Ψ^I that we generate using these inequalities rejects values of ψ_{EU10} lower than 0.416, but don't provide an upper bound. Ex-ante, we expect these inequalities to be more effective in finding a lower bound on the probability of an idiosyncratic delay ψ. To see why, assume that the expected revenue conditional on playing the optimal strategy are monotonically decreasing in ψ (i.e $\frac{\partial \tilde{V}(\mathcal{A}^\star(\psi),\psi;\cdot)}{\partial \psi} < 0$).[56] Then, it is impossible to find an alternative strategy A' such that

$$\tilde{V}_t\left(A', \psi'; S_{lt-1}, S_{-lt-1}\right) > \tilde{V}_t\left(\mathcal{A}^\star_{lt}\left(\psi^0\right), \psi^0; S_{lt-1}, S_{-lt-1}\right)$$

for $\psi' > \psi^0$.[57] It will still be possible to find A' such that $V\left(A', \psi'\right) > V\left(A^\star\left(\psi^0\right), \psi^0\right)$ for $\psi' < \psi^0$.[58] Our empirical results confirm this intuition: we find that our moment inequalities only provide a lower bound on the probability of an idiosyncratic delay.

7 Counterfactuals and Policy Analysis

7.1 *Delays in the absence of ERP*

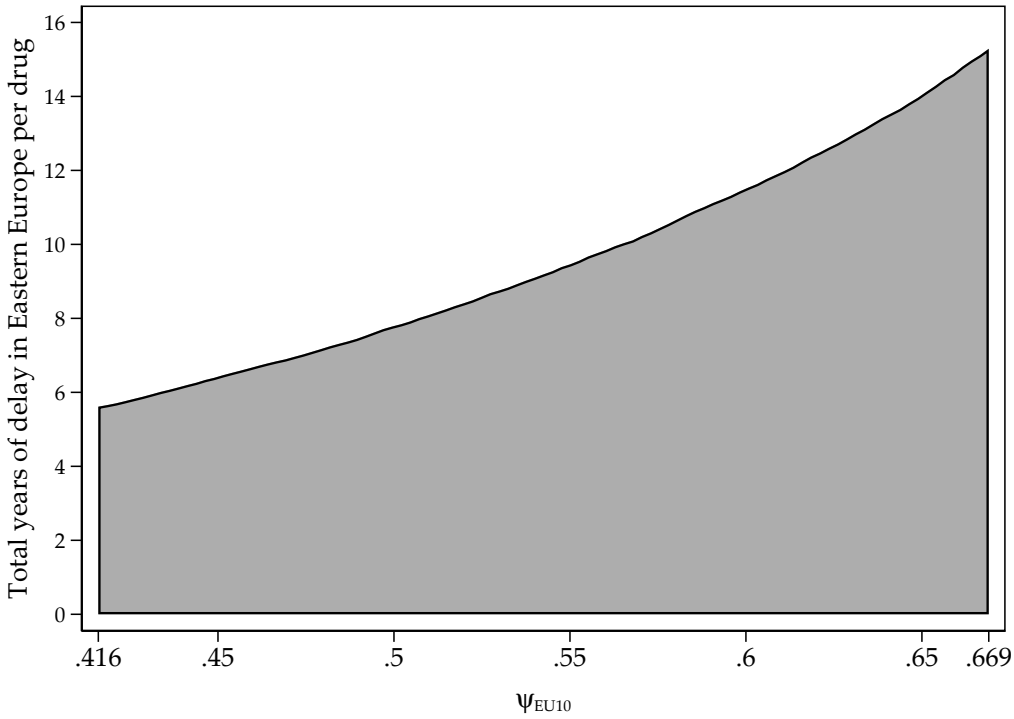
Using our estimates, we simulate how delays would change in the absence of ERP. In this scenario, firms no longer have any incentive to delay entry in any country, so he optimal strategy is to apply for entry everywhere right away. In this scenario, the only delays arise through the stochastic shock channel. Notice that this counterfactual exercise does not depend on the

[56] We do not have a formal proof of this statement. However, this is intuitively true: firms should be better off when the probability of a delay is lower, as they have better control over which entry sequence will be realized. For example, if the probability of delay were 0, the firm would be able to choose the profit-maximizing entry strategy. An increase in the probability of delay would reduce the likelihood of achieving the profit-maximizing entry sequence, therefore expected revenue would fall.

[57] To see why, notice that by assumption $V\left(A^\star\left(\psi^0\right), \psi^0\right) > V\left(A^\star\left(\psi'\right), \psi'\right)$ for all $\psi' > \psi^0$. Moreover, by definition of $A^\star(\cdot)$, $V\left(A^\star\left(\psi'\right), \psi'\right) > V\left(A', \psi'\right)$ for all ψ'.

[58] We could obtain an upper bound from data on expected revenue by showing that for a given value of ψ, no strategy will ever yield expected revenue as high as what the firm obtained in the data. The problem is that the firm achieves the highest possible expected revenue when playing the optimal strategy, and the reason we resort to moment inequalities is precisely that we cannot compute this strategy. However, we can find functions of data and parameters that always returns a value greater than the expected revenue under the optimal strategy. We describe this approach in the Online Appendix, including an example of a function that satisfies these requirements, but that unfortunately only yields an upper bound that is already ruled out from the moment inequalities that use entry data.

Figure 9: RANGE OF POTENTIAL DELAYS IN THE ABSENCE OF ERP



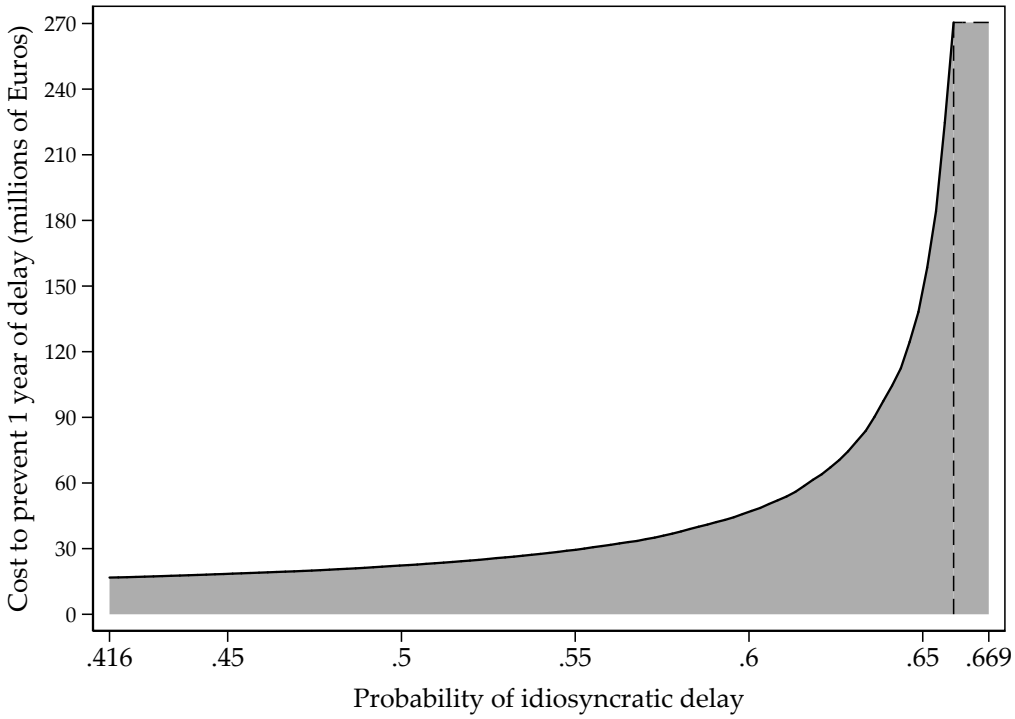
In this figure we plot the range of possible delays in Eastern European countries for the identified set of possible delay parameters. The y-axis reports the total delay of the average drug among all Eastern European countries. The upper bound of the interval returns the same average delay that is observed in the data (approximately 16 years, or 2 years per country), because it is calculated to match the overall probability of a delay.

specific pricing rule adopted to replace ERP, as long as it eliminates spillover effects across countries (this would also involve imposing limits on parallel trade, for example by carving out an exclusion for patent-protected pharmaceuticals). Importantly this also implies that we do not need to rely on our pricing model to work out the implications of this counterfactual. This is crucial as it is quite likely that such a large change in the regulatory environment would also bring about changes in the level and variation of prices across countries that our model might be unable to predict accurately.[59]

Figure 9 shows the range of possible delays in Eastern Europe implied by our identified set $[0.416, 0.669]$ for ψ_{EU10}. The parameter at the upper bound of the interval implies that idiosyncratic delays match observed delays. Under this scenario, the average drug should expect around 2 years of delay in each Eastern European country. At the lower bound of the interval, firms would instead expect an average delay of around 8-9 months in these countries, a reduction of approximately 63% in delays. Another potentially useful comparison is what would happen if we assigned to Eastern European countries the same delay parameter as the slowest Western European country, which is Italy. Italy has a delay probability of 0.590. If ψ_{EU10} took on this value, delays in the absence of ERP would fall by 27%, or around 6 months

[59]Many such policies have been suggested as a way to circumvent the effect of reference pricing. For example Towse et al. (2015) suggest that various forms of tiered pricing could be implemented centrally by the European Union to avoid delays in low-income countries. They propose Ramsey pricing and Value-Based Pricing as possible candidates.

Figure 10: COST OF REDUCING DELAYS BY 1 DRUG-YEAR IN EASTERN EUROPEAN COUNTRIES



In this figure we plot the range estimate of the cost of reducing delays by 1 drug-year in all 8 Eastern European countries in our sample. The y-axis is truncated above because the curve tends to infinity as ψ_{EU10} approaches 0.669, which is the value for which all delays are idiosyncratic, and therefore subsidies cannot change the delay profile.

in each country.

7.2 *Policy Analysis: implications of the impact of ERP on revenue*

Using our model, we can also calculate how much money firms would lose by disregarding the negative impact of reference pricing and applying for entry in all countries right away.[60] To get this number we simulate expected revenue for each drug under the naive strategy of applying everywhere right away and compare it to observed revenue. Since observed revenue depends on the realization of the delay shocks, we average across all drugs to obtain an estimate of the average loss across all drugs.[61]

We estimate this loss to be approximately €20 million. This is not a small sum, but it represents a small fraction of the average lifetime expected revenue for the drugs in our sample. The reason for this small number depends on two factors. First, prices in several Western European countries are only marginally higher than in Eastern Europe. This reduces the impact of reference pricing when it is applied. Second, our estimates suggest that countries with lower prices also have a higher probability of stochastic delays. Hence, even when firms launch ev-

[60] Notice that this is different than estimating how much revenue firms lose because of ERP, which would ultimately depend on what policy is chosen to replace ERP.

[61] The distribution of drug revenue is highly skewed, which means that the average is not necessarily the most informative moment that we could look at. However, given our data, we can only look at the average (instead of, say, the median) because we need to aggregate across drugs in order to remove the residual error that comes from the random realization of the delay shocks in the data.

erywhere at the same time, drugs tend to enter later in these countries, which also contributes to reduce the impact of ERP.

This result suggests that instead of removing ERP, the European Union could compensate firms in exchange for forgoing strategic delays. This solution has the advantage that it does not require EU Member States to give up the prerogative to manage drug pricing independently (which stems directly from the Treaty of Lisbon).[62] Subsidies could be handed out by a centralized European agency upon confirmation that an entry application has been sent and approved for entry in all European countries. The overall budget impact of this policy would be small according to our estimates. On average, during the period between 1995 and 2017, around 27 new drugs received approval in the EEA. Hence, the overall impact of this subsidy would be around €550 million per year, which would represent a very small fraction of the overall budget of the EU, which in 2016 was around €150 billion. It is also worth noting that from the point of view of a social planner, the lump sum constitutes a transfer, so any gains from early access, however small, would improve the overall welfare in the system.[63]

Figure 10 plots a distribution of the cost of reducing delays by 1 country-year for all countries in Eastern Europe over the range of the identified set of ψ_{EU10}. The purpose of this figure is not to argue whether paying the subsidy is worth it, but rather to illustrate the tradeoff. A larger probability of idiosyncratic delays implies that subsidies have a smaller impact on overall delays, hence the cost increases. At the upper bound, all delays are idiosyncratic, and therefore subsidies cannot affect the entry sequence of any drug. However, for more reasonable values of the parameter, the cost to increase access is not prohibitive. If we again assumed that ψ_{EU10} is equal to ψ_{ITALY}, then it would cost approximately €43 million to reduce delays by one drug-year in all Eastern European countries. Given that roughly 80 million people live in Eastern Europe, this comes at the cost of 50 cents per person.

8 Conclusion and Next Steps

This paper studies the extent to which external reference pricing policies contribute to the disparity in access to prescription drugs across countries. ERP generates complex incentives for firms who might benefit from strategically delaying entry in countries with low willingness to pay for drugs. Using a novel moment inequality approach, we characterize the impact of these policies on launch delays and on firm revenue. Our methodology allows us to obtain identification even though the firm's actions are unobserved, thus contributing to a growing body of literature showing how moment inequalities can make even the most complicated models tractable.

We can simulate one important scenario: the removal of reference pricing. In this case, the solution to the problem is easy to derive: firms no longer have any reason to delay entry in

[62] Article 168 of the Treaty on the Functioning of the European Union (i.e. the Treaty of Lisbon) explicitly states that "Union action in the field of public health shall fully respect the responsibilities of the Member States for the definition of their health policy and for the organization and delivery of health services and medical care and the allocation of the resources assigned to them."

[63] This conclusion might not hold in a model where additional frictions or dynamic considerations exist (e.g a shadow cost of raising governments funds, or dynamic implications on the incentives to invest in R&D). However, the size of the subsidy is small enough that these additional considerations will likely be second-order.

any country, so the optimal strategy is simply to send applications everywhere right away. We find that if ERP were eliminated delays in Eastern Europe would fall by up to 14 months per drug in each Eastern European country.

Policymakers and industry insiders are both aware of the externality generated by ERP, and several proposals to replace reference pricing with alternative systems have been suggested (Kanavos et al., 2011; OECD Health Policy Studies, 2008; Towse et al., 2015; Vogler et al., 2015). Our hope is that the analysis in this paper will contribute to the policy discussion on this important topic. One of our key results is that while the impact on delays is potentially large, revenue implications in the current equilibrium are relatively small. This suggests that firms could receive compensation and, in exchange, forgo strategic delays.

Our analysis also has some limitations. First, our model does not include a source of structural error, such as drug-country-year specific shocks in demand or price that are observed by the firm, but not by the econometrician. As we discuss in Section 5.1, unobservable demand shocks would likely lead us to underestimate the impact of strategic delays. Hence, this omission may lead us to underestimate the impact of reference pricing.

Second, we assume that our price model predicts exact prices. Introducing a structural error in the price estimation would create an insurmountable econometric challenge, since this error would propagate through reference pricing channel in ways that would be difficult to account for. One possible avenue to relax this assumption might be to assume that the structural error on price is known to the firm, but not to foreign governments. This would prevent the error from propagating across countries. However, the intuition behind the model would not change much. This type of error could help justify earlier-than-predicted entry, but the opposite (later-than-predicted entry) is usually much more common in the data.

Third, we assume that firms act as single agents even though our demand and price models imply that the actions of other firms can have an impact on revenue. Our demand model implies that entry of a competitor will negatively affect the market shares of all other products within the therapeutic class, and our price function suggests that entry of a competitor may have a negative impact on price. This introduces a certain degree of internal inconsistency in the model, though empirically, these effects may not be large enough to elicit a strategic reaction — in particular, the effect of competitors on price is very noisy according to our estimation results. Allowing for multiple agents is unfortunately not feasible in the current environment, as the expected revenue of alternative strategies in the moment inequality can only be computed holding the strategy of other firms fixed. This is a more general problem in the moment inequality literature, which has usually been more successfully applied to single-agent problems, rather than games involving multiple agents.

Finally, there are also elements that we do not model explicitly. We take the regulatory environment as exogenous. Presumably, the government could choose the reference pricing function according to some optimal decision-making criterion. We ignore this possibility here. The question of why reference pricing is effective in lowering prices is also interesting. Experience suggests that governments do not need to resort to ERP if they want to implement price cuts (we observe several instances of temporary price cuts in the data that do not seem to be related to reference pricing). Indeed, there is some evidence that reference pricing is some-

times used only as a pretense for cuts that would have occurred anyway. For example, Greece changed its reference pricing function in 2010 to a formula that resulted in roughly 10-20% lower prices across the board. It is likely that these cuts would have been mandated regardless due to the financial situation of the country at the time. These considerations go beyond the scope of the paper (we believe it is reasonable to assume that these changes are exogenous from the point of view of the firm) but might provide fertile ground for future research.

REFERENCES

Aitken, Murray, Ernst R. Berndt, D. Cutler, Michael Kleinrock, and Luca Maini, "Has The Era Of Slow Growth For Prescription Drug Spending Ended?," *Health Affairs*, sep 2016, *35* (9), 1595–1603.

Barnieh, Lianne, Fiona Clement, Anthony Harris, Marja Blom, Cam Donaldson, Scott Klarenbach, Don Husereau, Diane Lorenzetti, and Braden Manns, "A Systematic Review of Cost-Sharing Strategies Used within Publicly-Funded Drug Plans in Member Countries of the Organisation for Economic Co-Operation and Development," *PLoS ONE*, mar 2014, *9* (3), e90434.

Berry, Steven and Peter Reiss, "Empirical Models of Entry and Market Structure," in "Handbook of Industrial Organization," Vol. 3 2007, pp. 1845–1886.

Birg, Laura, "External Reference Pricing and the Choice of Country Baskets and Pricing Rules," *SSRN Electronic Journal*, 2016.

Borja, Garcia Lorenzo, "Modeling Global Pricing and Launching of New Drugs," 2014.

Brekke, Kurt R., Chiara Canta, and Odd Rune Straume, "Reference pricing with endogenous generic entry," *Journal of Health Economics*, 2016, *50*, 312–329.

_ , Ingrid Königbauer, and Odd Rune Straume, "Reference pricing of pharmaceuticals," *Journal of Health Economics*, 2007, *26* (3), 613–642.

Brekke, Kurt Richard, Tor Helge Holmas, and Odd Rune Straume, "Price Regulation and Parallel Imports of Pharmaceuticals," *Journal of Public Economics*, 2015, *129*, 92–105.

Bresnahan, Timothy F. and Peter C. Reiss, "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 1991, *99* (5), 977–1009.

Brown, Lawrence D., T. Tony Cai, and Anirban DasGupta, "Interval Estimation for a Binomial Proportion (with Discussion)," *Statistical Science*, 2001, *16* (2), 101–133.

Cardell, N. Scott, "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory*, 1997, *13* (02), 185.

Carone, Giuseppe, Christoph Schwierz, and Ana Xavier, *Cost-containment policies in public pharmaceutical spending in the EU* number September 2012.

Chaudhuri, Shubham, Pinelopi K. Goldberg, and Panle Jia, "The Effects of Global Patent Protection in Estimating Pharmaceuticals : A Case Study of Quinolones in India," *American Economic Review*, 2006, *96* (5), 1477–1514.

Cockburn, Iain M., Jean O. Lanjouw, and Mark Schankerman, "Patents and the global diffusion of new drugs," *American Economic Review*, 2016, *106* (1), 136–164.

Congressional Budget Office, "Prices for Brand-Name Drugs Under Selected Federal Programs," 2005, (June).

Costinot, Arnaud, Dave Donaldson, Margaret Kyle, and Heidi Williams, "The More We Die, The More We Sell? A Simple Test of the Home-Market Effect," 2016.

Council of European Communities, "Council Directive of 21 December 1988 relating to the transparency of measures regulating the pricing of medicinal products for human use and their inclusion in the scope of national health insurance systems," *Official Journal of the European Communities*, 1988, *L* (40), 8–11.

Crawford, Gregory S. and Matthew Shum, "Uncertainty and learning in pharmaceutical demand," *Econometrica*, 2005, *73* (4), 1137–1173.

Danzon, Patricia M., "Prices And Availability Of Pharmaceuticals: Evidence From Nine Countries," *Health Affairs*, oct 2003, pp. 521–536.

_ **and Andrew J. Epstein**, "Effects of regulation on drug launch and pricing in interdependent markets," *Advances in Health Economics & Health Services Research*, 2012, *23*, 35–71.

_ **and Michael F. Furukawa**, "Prices And Availability Of Biopharmaceuticals: An International Comparison," *Health Affairs*, sep 2006, *25* (5), 1353–1362.

_ **, Y. Richard Wang, and Liang Wang**, "The impact of price regulation on the launch delay of new drugs?evidence from twenty-five major markets in the 1990s," *Health Economics*, mar 2005, *14* (3), 269–292.

Decarolis, Francesco, "Medicare Part D: Are Insurers Gaming the Low Income Subsidy Design?," *American Economic Review*, apr 2015, *105* (4), 1547–1580.

Dickstein, Michael and Eduardo Morales, "What do Exporters Know?," 2018.

Drummond, Michael F., "Will there ever be a European drug pricing and reimbursement agency?," *The European Journal of Health Economics*, 2003, *4* (2), 67–69.

Dubois, Pierre, Ashvin Gandhi, and Shoshana Vasserman, "Bargaining and International Reference Pricing in the Pharmaceutical Industry," 2018.

Duggan, Mark and Fiona M. Scott Morton, "The Distortionary Effects of Government Procurement: Evidence from Medicaid Prescription Drug Purchasing," *Quarterly Journal of Economics*, 2006, *121* (1), 1–30.

Dunn, Abe, "Drug innovations and welfare measures computed from market demand: The case of anti-cholesterol drugs," *American Economic Journal: Applied Economics*, 2012, *4* (3), 167–189.

Duso, Tomaso, Annika Herr, and Moritz Suppliet, "The Welfare Impact of Parallel Imports: a Structural Approack Applied to the German Market for Oral Anti-Diabetics," *Health Economics*, 2014, *23* (9), 1036–1057.

Eizenberg, Alon, "Upstream innovation and product variety in the U.S. home PC market," *Review of Economic Studies*, 2014, *81* (3), 1003–1045.

European Commission, "Impact Assessment Accompanying the Document Proposal for a Directive of the European Parliament and of the Council relating to the transparency of measures regulating the price of medicinal products for human use and their inclusion in the scope of public," Technical Report 2012.

European Federation of Pharmaceutical Industries and Associations, "Principles for application of international reference pricing systems," Technical Report 2014.

European Parliament, "Directive 2001/83/EC of the European Parliament and of the Council," *Official Journal of the European Communities*, 2001, *L* (311), 67–128.

_ , "Directive 2004/27/EC, of 31 March 2004 amending Directive 2001/83/EC on the Community code relating to medicinal products for human use," *Official Journal of the European Union*, 2004, *L* (726), 34–57.

Grennan, Matthew, "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices," *American Economic Review*, 2013, *103* (1), 145–177.

_ **and Ashley Swanson**, "Transparency and Negotiated Prices: The Value of Information in Hospital-Supplier Bargaining," Technical Report, National Bureau of Economic Research 2016.

Holmes, Thomas J., "The Diffusion of Wal-Mart and Economies of Density," *Econometrica*, 2011, *79* (1), 253–302.

Houy, Nicolas and Izabela Jelovac, "Drug Launch Timing and International Reference Pricing," *Health Economics*, 2015, *24* (8), 978–989.

Illanes, Gastón, "Switching Costs in Pension Plan Choice," 2016.

Jaffe, Sonia and Mark Shepard, "Price-Linked Subsidies and Health Insurance Markups," 2017.

Kanavos, Panos, Sotiris Vandoros, Rachel Irwin, Elena Nicod, and Margaret Casson, "Differences in costs of and access to pharmaceutical products in the EU," *Policy Department A - Economic and Scientific Policy*, 2011.

Katz, Michael, "Essays in Industrial Organization and Technological Change." PhD dissertation 2007.

Kyle, Margaret K., "The role of firm characteristics in pharmaceutical product launches," *The RAND Journal of Economics*, sep 2006, *37* (3), 602–618.

_ , "Pharmaceutical Price Controls and Entry Strategies," *Review of Economics and Statistics*, 2007, *89* (February), 88–99.

_ **and Yi Qian**, "Intellectual Property Rights and Access to Innovation: Evidence from TRIPS," 2013.

Leopold, Christine, Aukje Katja Mantel-Teeuwisse, Leonhard Seyfang, Sabine Vogler, Kees de Joncheere, Richard Ogilvie Laing, and Hubert Leufkens, "Impact of External Price Referencing on Medicine Prices - A Price Comparison Among 14 European Countries," *Southern Med Review*, dec 2012, *5* (2), 34–41.

_ **, Sabine Vogler, A.K. Mantel-Teeuwisse, Kees de Joncheere, H.G.M. Leufkens, and Richard Laing**, "Differences in external price referencing in Europe - A descriptive overview," *Health Policy*, jan 2012, *104* (1), 50–60.

Manski, Charles F., *Partial Identification of Probability Distributions* Springer Series in Statistics, New York: Springer-Verlag, 2003.

Matteucci, Giorgio and Pierfrancesco Reverberi, "Drug innovation, price controls, and parallel trade," *International Journal of Health Economics and Management*, jun 2017, *17* (2), 159–179.

Morales, Eduardo, Gloria Sheu, and Andrés Zahler, "Extended Gravity," 2017.

OECD Health Policy Studies, *Pharmaceutical Pricing Policies in a Global Market* 2008.

Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii, "Moment Inequalities and Their Application," *Econometrica*, 2015, *83* (1), 315–334.

Panteli, Dimitra, Francis Arickx, Irina Cleemput, Guillaume Dedet, Helene Eckhardt, Emer Fogarty, Sophie Gerkens, Cornelia Henschke, Jennifer Hislop, Claudio Jommi, Daphne Kaitelidou, Pawel KawalecIlmo Keskimäki, Madelon Kroneman, Julio Lopez Bastida, Pedro Pita Barros, Joakim Ramsberg, Peter Vuorenkoski, Susan Spillane, Sabine Vogler, Lauri Vuorenkoski, Olivier Wouters, Helle Wallach Kildemoes, and Reinhard Busse, "Pharmaceutical regulation in 15 European countries," *Health Systems in Transition*, 2016, *18* (5), 1–118.

Pharmaceutical Industry Competitiveness Task Force, "Competitiveness and Performance Indicators 2005," Technical Report 2006.

Reich, Michael R., "The global drug gap," *Science*, 2000, *287* (March), 1979–1981.

Seim, Katja, "An empirical model of firm entry with endogenous product-type choices," *RAND Journal of Economics*, 2006, *37* (3), 619–640.

Stargardt, Tom and Jonas Schreyögg, "Impact of cross-reference pricing on pharmaceutical prices: Manufacturers' pricing strategies and price regulation," *Applied Health Economics and Health Policy*, 2006, *5* (4), 235–247.

Stern, Scott, "Market Definition and the Returns to Innovation: Substitution Patterns in Pharmaceutical Markets," 1996.

Thomson, Sarah and Elias Mossialos, "Primary care and prescription drugs: coverage, cost-sharing, and financial protection in six European countries.," *Issue brief (Commonwealth Fund)*, mar 2010, *82* (March), 1–14.

Toumi, Mondher, Cecile Rémuzat, Anne-Lise Vataire, and Duccio Urbinati, "External reference pricing of medicinal products : simulation- based considerations for cross- country coordination," Technical Report 2013.

Towse, Adrian, Michele Pistollato, Jorge Mestre-Ferrandiz, Zeba Khan, Satyin Kaura, and Louis Garrison, "European Union Pharmaceutical Markets: A Case for Differential Pricing?," *International Journal of the Economics of Business*, 2015, *22* (2), 263–275.

Vogler, Sabine, Lena Lepuschütz, Peter Schneider, and Verena Stühlinger, "Study on enhanced cross-country coordination in the area of pharmaceutical product pricing - Final Report," Technical Report Lot 2 2015.

Wilsdon, Tim, Eva Fiz, and Hugh Kirkpatrick, "The international impact of Swiss drug regulation," Technical Report 2013.

Young, K. E., I. Soussi, and M. Toumi, "The perverse impact of external reference pricing (ERP): a comparison of orphan drugs affordability in 12 European countries. A call for policy change," *Journal of Market Access & Health Policy*, 2017, *5* (1), 1369817.

Appendices

A Theoretical Derivations

In this Appendix we derive some results from the main body of the paper, and present a few complementary results that are not crucial to the results of the paper, but add robustness to the framework we use.

A.1 Logit Model

The utility of consumer ℓ, in country j, from consuming drug i (molecule m), belonging to therapeutic class κ, in year t is given by

$$u_{i(m,\kappa)\ell(j)t} = \delta_{ijt} + (\zeta_{m,\kappa} + (1-\sigma_\kappa)\,\varepsilon_{i\ell t})$$

where $\zeta_{m,\kappa}$, is common for all $i \in m$, and distributed according to the unique distribution such that if $\varepsilon_{i\ell t}$ is an extreme value random variable, then so is $\zeta_g + (1-\sigma)\,\varepsilon_{i\ell t}$ (Cardell, 1997). δ_{ijt} is parametrized as in Equation 5.

With this setup, one can show that the country j market share of *i within subset set m* is given by

$$s^m_{ijt} = \frac{\exp\left(\frac{\delta_{ijt}}{1-\sigma_\kappa}\right)}{D_m\left(\mathbf{X}_{mt}\right)} \tag{A.1}$$

where

$$D_m\left(\mathbf{X}_{mt}\right) = \sum_{k\in m} \exp\left(\frac{\delta_{kjt}}{1-\sigma_\kappa}\right)$$

and the market share of set *m within the overall market* is given by

$$s_{m/jt} = \frac{D_m\left(\mathbf{X}_{mt}\right)^{1-\sigma_\kappa}}{1 + \sum\limits_{h\in G_\kappa} D_h\left(\mathbf{X}_{ht}\right)^{(1-\sigma_\kappa)}} \tag{A.2}$$

where G_κ is the set of all molecules in class κ. Hence, the *overall market share of drug i* is

$$s_{ijt} = \frac{\exp\left(\frac{\delta_{ijt}}{1-\sigma_\kappa}\right) D_m\left(\mathbf{X}_{mt}\right)^{-\sigma_\kappa}}{1 + \sum\limits_{h\in G_\kappa} D_h\left(\mathbf{X}_{ht}\right)^{(1-\sigma_\kappa)}} \tag{A.3}$$

Derivation of the estimating equation Notice that the share of the outside option can be expressed as

$$s_{0jt} = \frac{1}{1 + \sum\limits_{h\in G_\kappa} D_h\left(\mathbf{X}_{ht}\right)^{(1-\sigma_\kappa)}} \tag{A.4}$$

Consider the log ratio of the market share of item i in group m to the outside good. According to the model, this can be expressed as

$$\ln\left(s_{ijt}\right) - \ln\left(s_{0jt}\right) = \left(\frac{\delta_{ijt}}{1-\sigma_\kappa}\right) - \sigma_\kappa\sigma \ln\left(D_m\left(\mathbf{X}_{mt}\right)\right)$$

Combining equations A.2 and A.4 we also obtain

$$\ln\left(D_m\left(\mathbf{X}_{mt}\right)\right) = \frac{\ln\left(s_{m/jt}\right) - \ln\left(s_{0jt}\right)}{1-\sigma_\kappa}$$

Hence we can write

$$\ln\left(s_{ijt}\right) - \ln\left(s_{0jt}\right) = \frac{\delta_{ijt}}{1-\sigma_\kappa} - \frac{\sigma_\kappa}{1-\sigma_\kappa\sigma}\left(\ln\left(s_{m/jt}\right) - \ln\left(s_{0jt}\right)\right)$$

which implies

$$\begin{aligned}
\left(1-\sigma_\kappa\right)\left(\ln\left(s_{ijt}\right) - \ln\left(s_{0jt}\right)\right) &= \delta_{ijt} - \sigma_\kappa\left(\ln\left(s_{m/jt}\right) - \ln\left(s_{0jt}\right)\right) \\
\implies \left(1-\sigma_\kappa\right)\ln\left(s_{ijt}\right) - \ln\left(s_{0jt}\right) &= \delta_{ijt} - \sigma_\kappa \ln\left(s_{m/jt}\right) \\
\implies \ln\left(s_{ijt}\right) - \ln\left(s_{0jt}\right) &= \delta_{ijt} + \sigma_\kappa \ln\left(\frac{s_{ijt}}{s_{m/jt}}\right)
\end{aligned} \tag{A.5}$$

A.2 Derivation of price estimating equation A.2

Let $p_{ijt}\left(\cdot\right)$ be defined as in equation 9. Then, for any $j, k \in \mathcal{N}_i$

$$\ln\left(\frac{p_{ijt}\left(\cdot\right)}{p_{ikt+1}\left(\cdot\right)}\right) = \begin{cases}
\ln\left(\frac{\gamma_j \cdot \exp\left(\beta_Z Z_{ijt} + \beta_D \ln\left(D_{ijt}\right)\right)}{\gamma_k \cdot \exp\left(\beta_Z Z_{ikt+1} + \beta_D \ln\left(D_{ikt+1}\right)\right)}\right) & \text{if } p^{\text{ref}}_{ijt}\left(\cdot\right) \geq p^{\text{gov}}_{ijt}\left(\cdot\right) \wedge \\
& p^{\text{ref}}_{ikt+1}\left(\cdot\right) \geq p^{\text{gov}}_{ikt+1}\left(\cdot\right) \\
\ln\left(\frac{\gamma_j \cdot \exp\left(\beta_Z Z_{ijt} + \beta_D \ln\left(D_{ijt}\right)\right)}{\left(1-\mu_k\right)\gamma_k \cdot \exp\left(\beta_Z Z_{ikt+1} + \beta_D \ln\left(D_{ikt+1}\right)\right) + \mu_k \tilde{p}^{\text{ref}}_{ikt+1}\left(\cdot\right)}\right) & \text{if } p^{\text{ref}}_{ijt}\left(\cdot\right) \geq p^{\text{gov}}_{ijt}\left(\cdot\right) \wedge \\
& p^{\text{ref}}_{ikt+1}\left(\cdot\right) < p^{\text{gov}}_{ikt+1}\left(\cdot\right) \\
\ln\left(\frac{\left(1-\mu_j\right)\gamma_j \cdot \exp\left(\beta_Z Z_{ijt} + \beta_D \ln\left(D_{ijt}\right)\right) + \mu_j \tilde{p}^{\text{ref}}_{ijt}\left(\cdot\right)}{\gamma_k \cdot \exp\left(\beta_Z Z_{ikt+1} + \beta_D \ln\left(D_{ikt+1}\right)\right)}\right) & \text{if } p^{\text{ref}}_{ijt}\left(\cdot\right) < p^{\text{gov}}_{ijt}\left(\cdot\right) \wedge \\
& p^{\text{ref}}_{ikt+1}\left(\cdot\right) \geq p^{\text{gov}}_{ikt+1}\left(\cdot\right) \\
\ln\left(\frac{\left(1-\mu_j\right)\gamma_j \cdot \exp\left(\beta_Z Z_{ijt} + \beta_D \ln\left(D_{ijt}\right)\right) + \mu_j \tilde{p}^{\text{ref}}_{ijt}\left(\cdot\right)}{\left(1-\mu_k\right)\gamma_k \cdot \exp\left(\beta_Z Z_{ikt+1} + \beta_D \ln\left(D_{ikt+1}\right)\right) + \mu_k \tilde{p}^{\text{ref}}_{ikt+1}\left(\cdot\right)}\right) & \text{if } p^{\text{ref}}_{ijt}\left(\cdot\right) < p^{\text{gov}}_{ijt}\left(\cdot\right) \wedge \\
& p^{\text{ref}}_{ikt+1}\left(\cdot\right) < p^{\text{gov}}_{ikt+1}\left(\cdot\right)
\end{cases} \tag{A.6}$$

where $\tilde{p}^{\text{ref}}_{ijt}\left(\cdot\right)$ is such that $p^{\text{ref}}_{ijt}\left(\cdot\right) = \tilde{p}^{\text{ref}}_{ijt}\left(\cdot\right) \cdot \theta_i$, and $\tilde{p}^{\text{ref}}_{ijt}\left(\cdot\right)$ is not a function of θ_i.

The proof of this Theorem hinges on showing that the reference price can be written as a linear function of the drug fixed effect. Intuitively, this result follows from the fact that the reference price can be written as a weighted average of government prices, which are all linear functions of the drug fixed effect. To prove the result of the theorem, we first prove the following Lemma:

Lemma 2. *Denote*

$$\lambda_{ijt} = \gamma_j \exp\left(\beta_Z Z_{ijt-1}\right)$$

and let $\lambda_{it} = \left\{\lambda_{ijt}\right\}_{j \in \mathcal{N}_i}$. *Then, there exists a set of weights* $\omega_{ijkt}\left(S_{t-1}, \{\lambda_{i\tau}\}_{\tau=1}^{t}\right)$ *such that for any drug* i, *country* j, *and year* t

$$p_{ijt}^{ref} = \sum_{\tau=1}^{t} \sum_{k \in R_{j\tau}} \omega_{ijkt}\left(S_{t-1}, \{\lambda_{i\tau}\}_{\tau=1}^{t}\right) p_{ik\tau-1}^{gov}\left(D_{ijt}\left(\xi_{jt}\right)\right)$$

Proof. This Lemma states that we can write p_{ijt}^{ref} as a linear function of government prices in all previous periods, using weights that depend only on the current entry sequences of all firms and structural parameters of the model other than the drug fixed effect. We use induction over t, where t indexes time starting with the year after the product was first approved (in the year of approval there cannot be any reference prices).

Consider $t = 1$. We want to show that for all j,

$$p_{ijt}^{\text{ref}} = \sum_{k \in R_{j1}} \omega_{ijk1}\left(S_0, \lambda_{i0}\right) p_{ik0}^{\text{gov}}\left(D_{ij1}\left(\xi_{j1}\right)\right) \tag{A.7}$$

The definition of the reference price is

$$p_{ij1}^{\text{ref}}\left(E_{i0}, D_{ij1}\left(\xi_{j1}\right)\right) = F_{j1}^{\text{ref}}\left(\left\{p_{ik0}\left(D_{ij1}\left(\xi_{j1}\right)\right)\right\}_{k \in \left(R_{j1} \cap E_{i0}\right)}\right)$$

Since reference pricing cannot be applied at time $t = 0$, the prices that can be referenced must be government prices:

$$p_{ik0}\left(D_{ij1}\left(\xi_{j1}\right)\right) = p_{ik0}^{\text{gov}}\left(D_{ij1}\left(\xi_{j1}\right)\right)$$

Then, if F_{j1}^{ref} is a linear function, like an average, equation A.7 is satisfied for constant weights. F_{j1}^{ref} can also be the minimum function, or the average of the three lowest prices, which are not linear.[64] These functions however, can be expressed as weighted averages, where the weights will depend on the relative ranking of the volume adjusted country government prices, which in turn will depend on λ_{ij0}.

In these cases weights can be constructed as follows. Recall that $E_{i0} = \left\{j : s_{j0} \neq 0\right\}$ is the set of countries where the product is available in period 0. This set can be obtained from information contained in S_0. We assume WLOG that E_{i0} is not empty (if it is, then there can be no possible reference and the case of $t = 1$ is identical to $t = 0$). Let n_k denote the rank of $k \in E_0$ in increasing order of λ_{i1}. In other words, if $n_k = 1$, then

$$\lambda_{ik1} = \min\left\{\lambda_{i\ell 1} : \ell \in \left(R_{j1} \cap E_0\right)\right\}$$

and, more generally,

$$\lambda_{ik1} = \min\left\{\lambda_{i\ell 1} : \ell \in \left(R_{j1} \cap E_0\right) \wedge n_\ell \geq n_k\right\}$$

[64] See Figure 1.

Hence, the country that ranks first is the one with the lowest government price, the one that ranks second has the second-lowest price, etc. Finally, let $m_{ij1} = \min\{|E_0|, 3\}$, where the operator $|\cdot|$ indicates the cardinality of a set.

If F_{j1}^{ref} is the average of the three lowest prices, the weights can be written as

$$\omega_{ijk1}(S_0, \lambda_{i0}) = \begin{cases} 1 & \text{if } n_k = 1 \\ 0 & \text{otherwise} \end{cases}$$

If F_{j1}^{ref} is the average of the three lowest prices instead, construct the weights as

$$\omega_{ijk1}(S_0, \lambda_{i0}) = \begin{cases} \frac{1}{m_{ij0}} & \text{if } n_k \leq m_{ij0} \\ 0 & \text{otherwise} \end{cases}$$

These weights are written as a function of S_0 and λ_{i0} only, hence they satisfy the premise of the proposition.

To conclude the proof, suppose that the assertion of the proposition is true for $\tau \in \{1, \ldots, t-1\}$, and show that it must hold for t as well. This is easy to prove. We can walk through the same exact steps as we did for $t = 1$, but substituting $p_{ikt-1}\left(E_{it-1}, D_{ijt}\left(\xi_{jt}\right)\right)$ for λ_{i1}. This will give us weights for p_{ijt}^{ref} as a linear function of the prices in the previous period. By construction, prices in the previous period are a weighted average of government prices and reference prices. The reference prices are themselves linear functions of adjusted government prices by the inductive assumption. Since the sum of linear functions is also linear, the proposition must hold for period t as well. ■

The Lemma gives us a way to write the reference price as a weighted average of government prices. Since government prices are a multiplicative function of θ_i, so are reference prices. Hence we can write

$$p_{ijt}^{\text{ref}}(\cdot) = \tilde{p}_{ijt}^{\text{ref}}(\cdot) \cdot \theta_i$$

where $\tilde{p}_{ijt}^{\text{ref}}$ is not a function of θ_i. Then, we can write

$$p_{ijt}(\cdot) = \begin{cases} \theta_i \gamma_j \cdot \exp\left(\beta_Z Z_{ijt} + \beta_D \ln\left(D_{ijt}\right)\right) & \text{if } p_{ijt}^{\text{ref}}(\cdot) \geq p_{ijt}^{\text{gov}}(\cdot) \\ \theta_i \gamma_j \cdot \exp\left(\beta_Z Z_{ijt} + \beta_D \ln\left(D_{ijt}\right)\right) + \mu_j \tilde{p}_{ijt}^{\text{ref}}(\cdot)\,\theta_i & \text{if } p_{ijt}^{\text{ref}}(\cdot) < p_{ijt}^{\text{gov}}(\cdot) \end{cases}$$

which shows that $p_{ijt}(\cdot)$ is a multiplicative function of θ_i. Hence, when we consider $\frac{p_{ijt}}{p_{ikt+1}}$, θ_i will appear in both the denominator and the numerator and we can eliminate it, giving us the result in equation A.6.

■

A.3 Foundations for the Price-Setting Equation

In the main body of the paper we do not provide a micro foundation of the price setting equation from a utility- or revenue-based optimization model. The characteristics of such a model

are less relevant for our paper. However, in this section we provide two possible sets of assumptions that could justify an estimation equation identical to the one we use.

Foundations as a static Nash Bargaining Model

We assume that the firm and the government play a Nash Bargaining game. The game is repeated every period, but for simplicity the two parties only split the static welfare gains from the current period. In reality, prices impose dynamic constraints through reference pricing that both agents should take into account. To eliminate these dynamic considerations one could assume that the government is a myopic agent and that the firm's bargaining unit is only tasked with carrying out the negotiation, without concerns for the future ramifications of the agreed-upon price.[65]

The equilibrium price in a standard Nash Bargaining model is given by

$$p^{\star} = \arg\max_{p} \left[\Delta W_{ijt}\right]^{b_j} \times \left[\Delta \Pi_{ijt}\right]^{1-b_j}$$

where ΔW_{ijt} represents the change in the welfare of the government from having drug i available, $\Delta \Pi_{ijt}$ represents the incremental change in revenue, and b_j is the bargaining power of country j. Notice that the interpretation of ΔW_{ijt} is not necessarily welfare, but could more generally be described as the objective function of the government agent tasked with completing the negotiation.

Under our assumptions of static bargaining, $\Delta \Pi_{ijt}$ is simply the potential revenue in country j, and since demand is price inelastic, we can divide through by demand to recast the problem as a negotiation over the unit price of the product (instead of total revenue). We abstract away from marginal costs of production since for brand drugs they are a negligible fraction of prices. The simplified problem can be written as

$$p^{\star} = \arg\max_{p} \left[\Delta w_{ijt} - p\right]^{b_j} \times \left[p\right]^{1-b_j}$$

where the interpretation of Δw_{ijt} is the average change in the welfare function from obtaining an additional unit of drug i. The standard Nash bargaining solution, can then be written as

$$p^{\star} = \Delta w_{ijt} \left(1 - b_j\right)$$

This price denotes the equilibrium in the absence of reference pricing, and represents the *government price* p_{ijt}^{gov}.

To account for the impact of reference pricing we propose that the government can negotiate more effectively by eliciting a signal about what prices are charged abroad. We incorporate this possibility in the model by assuming that the signal (i.e. the reference price) affects the

[65] There are several reasons why short-term considerations could in fact play a major role for most government agencies. First, the main goal of pharmaceutical agencies is to keep spending within the limits of their budget, which is often specifically carved out for prescription drugs thus limiting the ability to generate trade-offs such as paying more for cost-effective drugs that would save money in other areas of health care, such as inpatient care. Turnover of government officials might also contribute to the failure of adopting long-term strategies. On the pharmaceutical company side, most firms have a separate bargaining unit for each country. Informal conversations with industry insiders seem to suggest that these unit operate in relative independence from one another.

bargaining weight assigned to the government. The reference price p_{ijt}^{ref} is calculated as described in the main body of the paper.[66] Given p_{ijt}^{ref}, we write the bargaining weight of the government as

$$B_{ij}\left(p_{ijt}^{\text{ref}}\right) = b_j + \left(1 - b_j\right)\mu_j \left(1 - \frac{p_{ijt}^{\text{ref}}}{p_{ijt}^{\text{gov}}}\right) \cdot \mathbb{I}\left\{p_{ijt}^{\text{ref}} < p_{ijt}^{\text{gov}}\right\}$$

where $p_{ijt}^{\text{gov}} = \Delta w_{ijt}\left(1 - b_j\right)$ is a function of model parameters that reflects the price that the government would have obtained without using reference pricing. We define the bargaining weight of the firm as $1 - B_{ij}\left(p_{ijt}^{\text{ref}}\right)$.

The function $B_{ij}\left(\cdot\right)$ has several attractive properties. First, it reduces to the base case whenever $p_{ijt}^{\text{ref}} < p_{ijt}^{\text{gov}}$. This has the intuitive implication that observing a reference price that is higher than the country's own internal benchmark does not affect negotiations. Second, the bargaining weight is inversely proportional to the reference price, meaning that a lower reference price lets the government extract a greater discount. Third, as long as $\mu_j \in (0,1)$ the bargaining weight is also lies on the unit interval, which insures an interior solution for the first-order condition.

The first-order condition of the Nash Bargaining problem with the specified bargaining weights is

$$\begin{aligned}[p]: \quad & \left(\Delta w_{ijt} - p\right)^{-1+b+\mu_j - b_j\mu_j - \frac{\mu_j p_{ijt}^{\text{ref}}}{\Delta w_{ijt}}} \cdot \\ & \left(\left(1 - b_j\right)\left(1 - \mu_j\right)\Delta w_{ijt} + \mu_j p_{ijt}^{\text{ref}} - \Delta w_{ijt}\right) p^{-b_j\left(1-\mu_j\right) + \mu_j\left(1 - \frac{p_{ijt}^{\text{ref}}}{\Delta w_{ijt}}\right)} = 0\end{aligned}$$

and has three roots:

$$\begin{aligned} p_1^\star &= 0 \\ p_2^\star &= \Delta w_{ijt} \\ p_3^\star &= \left(1 - \mu_j\right)\left(1 - b_j\right)\Delta w_{ijt} + \mu_j p_{ijt}^{\text{ref}} \end{aligned}$$

Notice that $p_{ijt}^{\text{ref}} \leq \left(1 - b_j\right)\Delta w_{ijt}$ whenever the reference price binds. Hence, $p_1^\star < p_3^\star < p_2^\star$.

The second-order condition is given by

$$\begin{aligned}SOC: \quad & p^{-1-b\left(1-\mu_j\right)-\mu_j\left(1 - \frac{p_{ijt}^{\text{ref}}}{\Delta w_{ijt}}\right)} \cdot \left(\Delta w_{ijt} - p\right)^{-2+b_j+\mu_j - b_j\mu_j - \frac{\mu_j p_{ijt}^{\text{ref}}}{\Delta w_{ijt}}} \cdot \\ & \left(\Delta w_{ijt}\left(1 - b_j\right)\left(1 - \mu_j\right) + \mu_j p_{ijt}^{\text{ref}}\right)\left(-b_j \Delta w_{ijt}\left(1 - \mu_j\right) + \mu_j\left(p_{ijt}^{\text{ref}} - \Delta w_{ijt}\right)\right)\end{aligned}$$

[66] As a side note, the reference price does not necessarily need to be linked to other prices, but can be anything else that might affect negotiations.

and, for $p \in (0, \Delta w_{ijt})$, is proportional to

$$\begin{aligned} SOC &\propto \left(\Delta w_{ijt}\left(1-b_j\right)\left(1-\mu_j\right)+\mu_j p_{ijt}^{\text{ref}}\right)\left(-b_j \Delta w_{ijt}\left(1-\mu_j\right)+\mu_j\left(p_{ijt}^{\text{ref}}-\Delta w_{ijt}\right)\right) \\ &\propto \left(-b_j \Delta w_{ijt}\left(1-\mu_j\right)+\mu_j\left(p_{ijt}^{\text{ref}}-\Delta w_{ijt}\right)\right)<0 \end{aligned}$$

Hence the objective function is maximized for $p = p_3^\star$. The two other roots of the first-order condition are also roots for the second-order condition, therefore they represent points of inflection.

The final solution to this bargaining problem is therefore made up of two equations:

$$p_{ijt} = \begin{cases} \left(1-b_j\right)\Delta w_{ijt} & \text{if } p_{ijt}^{\text{ref}} \geq \left(1-b_j\right)\Delta w_{ijt} \\ \left(1-\mu_j\right)\left(1-b_j\right)\Delta w_{ijt}+\mu_j p_{ijt}^{\text{ref}} & \text{if } p_{ijt}^{\text{ref}} < \left(1-b_j\right)\Delta w_{ijt} \end{cases}$$

This solution will have the same form of our estimating equation as long as $\left(1-b_j\right)\Delta w_{ijt}$ can be written as a function of the observables we have included in our parametric function for the government price.

A.4 Proof of Theorem 1

To prove the theorem we will rely on the strong law of large numbers applied to non-identical, independent random variables. For this reason it will be useful to define the payoff of firm l as a random variable. We will then show that the set of these random variables (one for each firm l) satisfies the Kolmogorov criterion, which in turn implies the strong law of large numbers.

Let $\tilde{\mathbf{V}}_{lt}\left(\mathcal{A}_{it}; S_{lt-1}, S_{-lt-1}\right)$ denote the payoff of the firm starting in period t, conditional on the value of state variable in period $t-1$ and on firm l following strategy $\mathcal{A}_{lt}$. Note that by definition, $\mathbb{E}\left[\tilde{\mathbf{V}}_{lt}\left(\mathcal{A}_{it}; S_{lt-1}, S_{-lt-1}\right)\right] = \tilde{V}_t\left(\mathcal{A}_{lt}; S_{lt-1}, S_{-lt-1}\right)$.

For the proof of the Theorem we need the following Lemma and Corollary.

Lemma 3. *Let $\Pi\left(S_l, S_{-l}\right)$ be defined as in equation 11. Then $\Pi\left(S_l, S_{-l}\right)$ is finite.*

Proof. We prove that $\Pi_\tau\left(S_l, S_{-l}\right)$ is finite by showing that period profits are bounded. The realization of period profits depends on ξ_{jt}. Define

$$\Pi_\tau\left(S_l, S_{-l}, \xi_{j\tau}\right) = \sum_{i \in \mathcal{I}_l} \sum_{j \in S_{i\tau}} p_{ij\tau}\left(S_{l\tau-1}, S_{-l\tau}, D_{ij\tau}\left(\xi_{j\tau}\right)\right) \cdot D_{ij\tau}\left(S_{-l\tau}, \xi_{j\tau}\right)$$

For any given product i and country j, we can write demand as

$$D_{ij\tau}\left(S_{-l\tau}, \xi_{j\tau}\right) = MS_{j\tau} \cdot \frac{\exp\left(\alpha_{ij} + \beta_i \text{age}_{i\tau} + \eta_i NF_{ij\tau} + \xi_{j\tau}\right)}{1 + \sum_{\ell \in E_{-l\tau}} \exp\left(\alpha_{\ell j} + \beta_\ell \text{age}_{\ell\tau} + \eta_\ell NF_{\ell j\tau} + \xi_{j\tau}\right)}$$

where $MS_{j\tau}$ is the market size in country j in period τ. Hence, $D_{ij\tau}\left(S_{-l\tau}, \xi_{jt}\right) \in \left(0, MS_{j\tau}\right)$.

Price is bounded above by the government price:

$$p_{ij\tau}\left(S_{l\tau-1}, S_{-l\tau}, D_{ij\tau}\left(\xi_{j\tau}\right)\right) \leq p_{ij\tau}^{\text{gov}}\left(S_{l\tau-1}, S_{-l\tau}, D_{ij\tau}\left(\xi_{j\tau}\right)\right)$$

Moreover, using the definition of government price in equation 7 we can rewrite

$$p_{ij\tau}^{\text{gov}}\left(S_{l\tau-1}, S_{-l\tau}, D_{ij\tau}\left(\xi_{j\tau}\right)\right) \cdot D_{ij\tau}\left(\xi_{j\tau}\right) = p_{ij\tau}^{\text{gov}}\left(S_{l\tau-1}, S_{-l\tau}, 1\right) \cdot \left(D_{ij\tau}\left(\xi_{j\tau}\right)\right)^{1+\beta_D}$$

Hence, the period payoff of a single drug in any given country is bounded above by

$$p_{ij\tau}^{\text{gov}}\left(S_{l\tau-1}, S_{-l\tau}, 1\right) \cdot \left(MS_{j\tau}\right)^{1+\beta_D}$$

and bounded below by 0. This implies that the period payoff in any given country and period is finite, and therefore $\Pi_\tau\left(S_l, S_{-l}\right)$ is also finite. ■

Corollary 4. $\tilde{\mathbf{V}}_{lt}\left(\mathcal{A}_{it}; S_{lt-1}, S_{-lt-1}\right)$ *has finite variance.*

Proof. This lemma follows directly from Lemma 3. $\tilde{\mathbf{V}}_{lt}\left(\mathcal{A}_{it}; S_{lt-1}, S_{-lt-1}\right)$ is defined as the discounted sum of the expected period payoffs. By Lemma 3 the expected period payoffs are finite. Let

$$\Pi^{UB} = \max_{\left(S_l, S_{-l}\right)} \Pi_\tau\left(S_l, S_{-l}\right)$$

Then, the support of $\tilde{\mathbf{V}}_{lt}\left(\mathcal{A}_{it}; S_{lt-1}, S_{-lt-1}\right)$ is bounded above by $(T-t) \cdot \Pi^{UB} < \infty$.

Since $\tilde{\mathbf{V}}_{lt}\left(\mathcal{A}_{it}; S_{lt-1}, S_{-lt-1}\right)$ is also bounded below by 0, it must have finite variance. ■

At this point we are ready to prove Theorem 1.

Proof of Theorem 1. For any given drug i, $V_t\left(S_l^o, S_{-l}^o\right)$ represents a draw from the distribution of $\tilde{\mathbf{V}}_{lt}\left(\mathcal{A}_{it}^\star; S_{lt-1}, S_{-lt-1}\right)$. By Corollary 4, $\text{Var}\left(\tilde{\mathbf{V}}_{lt}\left(\mathcal{A}_{it}^\star; S_{lt-1}, S_{-lt-1}\right)\right) < \infty$. Moreover, the random variables $\tilde{\mathbf{V}}_{lt}\left(\mathcal{A}_{it}^\star; S_{lt-1}, S_{-lt-1}\right)$ are independently distributed. Thus, our premise satisfies the Kolmogorov criterion, which implies that the strong law of large numbers applies to the sequence of random variables $\tilde{\mathbf{V}}_{lt}\left(\mathcal{A}_{it}^\star; S_{lt-1}, S_{-lt-1}\right)$, and the sample average of the realized payoffs will converge to the average of their expected values. Formally, for any $\varepsilon > 0$, we can find M' such that

$$\frac{1}{M}\sum_{i=1}^{M}\left(V_t\left(S_l^o, S_{-l}^o\right) - \tilde{V}_t\left(\mathcal{A}_{lt}^\star; S_{lt-1}, S_{-lt-1}\right)\right) < \varepsilon$$

for all $M > M'$. This concludes the proof.

■