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

## Online Appendix

Luca Maini* Fabio Pammolli†

January 31, 2022

## Contents



*UNC Chapel Hill, email: lmaini@email.unc.edu
†Politecnicao di Milano, email: fabio.pammolli@polimi.it

This Online Appendix has four sections. The first section describes the data and the process of creating the dataset used to estimate the structural model. The second section provides additional background on the European pharmaceutical market. The third section presents additional empirical results and robustness checks that were not included in the main body of the paper. The fourth section presents theoretical results that complement and extend the model presented in the paper.

# A Data Appendix

## A.1 Description of the MIDAS dataset

### A.1.1 Data coverage

The IMS MIDAS database we use has data on retail (i.e. pharmacies) and hospital sales for 27 European countries and 48 quarters, starting in 2001 (either Q2 or Q3) and ending in 2013 (either Q1 or Q2), giving us almost complete coverage of the European Economic Area during that time span.[1]

Several countries are partially missing data. Most commonly this means missing certain years of data, missing hospital sales data, and having sales data in Euro instead of local currency (for countries where the Euro is not the local currency). The problem with having data in Euros instead of the local currency is that the raw data were originally converted using a fixed exchange rate (for the entire 12-year sample) instead of a quarter-specific exchange rate. This affects makes the sales figures unreliable in countries that do not have a fixed exchange rate with the Euro. Table 1 describes our data coverage by country.

Since the problem that we study concerns spillover effects across countries, it is imperative that we try to include in our sample as many countries as possible. Ultimately, we end up discarding 2 of the 27 countries in our sample due to unsolvable data problems: Czech Republic and Slovakia. These two countries had non-Euro currencies with fluctuating exchange rates during our sample period, which means that their sales data are unreliable. We include them in the reduced form results that do not involve prices but drop them from the structural analysis.

Twelve of the remaining 25 countries have some missing data: Denmark, Estonia, Greece, Ireland, Latvia, Lithuania, Luxembourg, Netherlands, Portugal, Romania, Slovenia, and Sweden. Some of these are minor issues that do not affect the analysis. Denmark and Slovenia report retail and hospital sales jointly, which limits our ability to separate retail and hospital products in some robustness checks but leaves our main specification unaffected. We have sales in Euros only for Denmark and Estonia, but the currency of these two countries was pegged to the Euro throughout the duration of our sample period, so the conversion leaves the sales figures unaffected. Slovenia also only reports sales in Euros, and for the first 3 years of our sample data its currency was not

[1] We are missing data for Cyprus, Iceland, Lichtenstein, and Malta. Croatia is also a current member of the EEA, but only entered in 2015, which is outside our sample.

Table 1: DATA AVAILABILITY BY COUNTRY

| Country | Years | Hospital Sales Data | Currency |
|---|---|---|---|
| Austria | 2001Q2-2013Q1 | Available | Euro |
| Belgium | 2001Q2-2013Q1 | Available | Euro |
| Bulgaria | 2001Q3-2013Q2 | Available | Local |
| Czech Republic | 2001Q2-2013Q1 | Available | Euro[a] |
| Denmark | 2001Q2-2013Q1 | Not distinguished separately | Euro[b] |
| Estonia | 2001Q3-2013Q2 | Not available | Euro[c] |
| Finland | 2001Q2-2013Q1 | Available | Euro |
| France | 2001Q2-2013Q1 | Available | Euro |
| Germany | 2001Q2-2013Q1 | Available | Euro |
| Greece | 2001Q2-2013Q1 | Not available | Euro |
| Hungary | 2001Q3-2013Q2 | Available | Local |
| Ireland | 2001Q2-2013Q1 | Available starting 2006Q1 | Euro |
| Italy | 2001Q2-2013Q1 | Available | Euro |
| Latvia | 2001Q3-2013Q2 | Not available | Local[d] |
| Lithuania | 2001Q3-2013Q2 | Available starting 2002Q3 | Local[d] |
| Luxembourg | 2001Q2-2013Q1 | Not available | Euro |
| Netherlands | 2007Q2-2013Q1 | Not available | Euro |
| Norway | 2001Q3-2013Q2 | Available | Local |
| Poland | 2001Q3-2013Q2 | Available | Local |
| Portugal | 2001Q2-2013Q1 | Available starting 2010Q1 | Euro |
| Romania | 2001Q3-2013Q2 | Available starting 2005Q1 | Local |
| Slovakia | 2001Q3-2013Q2 | Available | Euro[e] |
| Slovenia | 2001Q3-2013Q2 | Not distinguished separately | Euro[f] |
| Spain | 2001Q2-2013Q1 | Available | Euro |
| Sweden | 2004Q1-2013Q2 | Available | Local |
| Switzerland | 2001Q3-2013Q2 | Available | Local |
| UK | 2001Q3-2013Q2 | Available | Local |

[a] The Czech koruna appreciated substantially against the Euro between 2001 and 2012. The average rate fell from 34.021 Ck for 1 Euro in 2002Q1, to 25.167 Ck for 1 Euro in 2012Q4.
[b] The value of a Danish Krona oscillated between 0.13398 and 0.13468 Euros in the period from 2001 to 2012.
[c] The Estonian kroon was pegged to the Euro (at a rate of 15.6466 krooni per Euro) until adoption of the Euro in 2011.
[d] Latvia and Lithuania adopted the Euro after the end of our data (in 2014 and 2015 respectively).
[e] The Slovak koruna appreciated substantially against the Euro until its adoption in 2009. The average rate fell from 42.234 Sk for 1 Euro in 2002Q1, to 30.35 Sk for 1 Euro in 2008Q4.
[f] The value of the Slovenian tolar depreciated from 219.3683 to 239.9533 SIT for 1 Euro between 2001Q3 and 2004Q3. After that, the exchange rate remained constant until the country adopted the Euro in 2007.

Table 2: VARIABLES IN THE IMS MIDAS DATABASE

| Variable Name | Description | Variable Type |
|---|---|---|
| country | Country | |
| distributionchannel | Distribution Channel | Hospital/Retail |
| mlist | Active Ingredient | Product definition |
| crp | Marketing Firm | Product definition |
| atc4 | ATC4 | Product definition |
| internationalproduct | International Product Name | Product definition |
| prd | Local Product Name | Product definition |
| productform | Product Form | Product definition |
| productinternationalstrength | Strength | Product definition |
| internationpack | International Package Identifier | Product definition |
| localpack | Local Package Identifier | Product definition |
| rxorotc | Rx or OTC | Product characteristics |
| localproductlaunchdate | Local Product Launch date | Product characteristics |
| estpatentexpirydate | Estimated Patent Expiration Date | Product characteristics |
| protection | Protection Status | Product characteristics |
| producttype | Product Type | Product characteristics |
| licensinginformation | Licensing Information | Product characteristics |
| sales_mnf_qtr_*'qrt'*_*'yr'*_local_curr | Sales in local currency | Sales information |
| sales_mnf_qtr_*'qrt'*_*'yr'*_lceuro | Sales in Euro | Sales information |
| standard_units_qtr_*'qrt'*_*'yr'* | Quantity sold | Sales information |

pegged to the Euro. However, the Slovenian Tolar fluctuated considerably less than the Czech and Slovak currencies, so we ultimately decided to keep the data in the sample (without additional corrections).

The remaining data issues concern the lack of hospital sales data in 9 countries, and missing years for Netherlands and Sweden. We address these problems by imputing the missing data. These steps are detailed in section A.4.

### A.1.2 Variables

The variables contained in each file are described in Table 2. Each drug is identified by a combination of active ingredient (molecule), marketing firm, therapeutic class (defined by the ATC4 classification), name (both international and country-specific), form, strength, and package.[2] Be-

[2]For more details on the ATC classification, please see http://www.ephmra.org/classification.

sides revenue and quantity sales, the data provide a few additional characteristics: whether the product is a prescription or an over-the-counter medication, the launch date, the estimated patent expiration date, the product type (brand or generic), protection status, and licensing information.

### A.1.3 Prices and rebates

The sales figures in the MIDAS database do not include hidden discounts and rebates that firms occasionally grant to private payers. While we acknowledge this issue, we note that estimates of non-US rebates tend to be low. For example, Feng (2020) shows in Appendix G that there is a near-perfect correspondence between US rebates and overall rebates in financial filings of pharmaceutical companies (Figure A1). However, his analysis is limited to the market for statins. Other markets may function in different ways. We also conducted several informal interviews with pharmaceutical executives with firsthand knowledge of pricing in Europe. All the people we talked to (all of whom asked to remain anonymous), confirmed that rebates for brand drugs in Europe exist, but are in a range from single-digit to low teens (in percentage terms).

Nonetheless, we believe it is important to consider how the presence of rebates might affect our analysis. Below, we discuss some possible scenarios. While some of them suggest that the presence of rebates may lead us to overestimate strategic delays, most scenarios actually imply that the presence of rebates makes our model more conservative.

The presence of rebates create two issues. First hidden rebates imply that we are overestimating revenues and prices, and disproportionately so in countries where rebates are higher. Second, governments may or may not observe rebates, which in turn affects our estimation of the reference pricing mechanism. We consider each issue separately.

Consider our estimation of revenue and prices. The presence of hidden rebates in lower income countries does not significantly affect the predictions of our model. To see why consider a concrete example in a stylized, two-country setting. Suppose the government of Poland decides to grant a hidden rebate to Pfizer in order to induce them to launch their drug sooner. By doing so they prevent the government of Italy from applying a costly reference. In this case, we would see the product enter Italy and Poland within a relatively short interval. Since the observed "list" price has to be high enough that Italy does not reference it (or they do, but the impact is very small), our model predicts that it makes sense for Pfizer to launch everywhere, and therefore there are no strategic delays. The only meaningful difference is that our model will predict that Pfizer's revenue in Poland is higher than it really is (because it does not account for the rebate). In general, the reason behind the absence of strategic delays does not matter. Poland may surrender a high price, or a high "list" price, with a hidden rebate. From the point of view of the model, the two are equivalent.

The presence of hidden rebates in higher income countries may be more problematic. Hidden rebates in higher-income countries raise the possibility that a lower price elsewhere might not actually matter for reference pricing. In the example above, suppose Pfizer sets a price of 10 in Italy, but with a 50% discount. Later, it enters Poland, at a price of 7. In our data, a few things

might happen:

1. The "list" price in Italy falls from 10 to 7, but the "real" price of 5 does not move. In this case, our model could wrongly attribute any idiosyncratic delay to entry in Poland to a strategic motive.

2. The list price in Italy does not move. In this case, our pricing equation may be affected, as we would underestimate the $\mu_k$ coefficient for Italy. This is a problem for our estimation, but, in general, lower values of $\mu_k$ lead to more conservative estimates of strategic delays, so, if anything, this may lead us to underestimate the extent of strategic delays.

More generally, the presence of hidden rebates in higher countries means we overestimate revenue in these countries, which in turn implies that we will overestimate the how important reference pricing is, and in turn could bias our estimates upward.

Next, consider the issue of whether or not governments observe rebates. If governments do not observe rebates, then our reference pricing function is actually specified correctly, because we have access to the same data that the government has access to. If governments do observe rebates, the true specification of the model would need to include a structural error in the price equation. As noted in the main body of the paper, this makes estimation incredibly challenging, as the unobserved error would propagate through the reference pricing mechanism in ways that seem hard to account for. It is unclear what impact this would have on our estimates. If we believed this to be the case—and that rebates were sizable—we probably would consider estimation of the reference pricing mechanism to be a futile exercise. However, we believe that this is the least likely scenario.

Our conclusion is that any assumption about imperfect information would be ad-hoc, so using the data as is seems to us the most reasonable course of action.

### A.2 Additional data sources

Additional data sources are reported in Table 3. We use approval dates from the EMA and the Heads of Medicines Agencies to calculate launch delays. We use data from the Global Burden of Disease study to calculate market size for the demand system. We use exchange rates to transform sales reported in local currencies to Euros. Finally, we use data on GDP and population to run some simple reduced-form checks.

All these additional data sources are publicly available (the links in Table A.2 are accurate as of November 20th, 2017). Updated instructions on how to find more recent versions of these data can be obtained in the README file of the replication package associated with this paper.

### A.3 Data cleaning and construction of the main dataset

The dataset we ultimately use in the analysis is a collapsed version of the original dataset described in section A.1. In this section we describe how we construct each of the main variables in that dataset.

Table 3: ADDITIONAL DATA SOURCES

| Data | Use | Source | Source Link |
|---|---|---|---|
| Drug Approval Date (EMA) | Calculate launch delays | European Medicines Agency | http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/landing/epar_search.jsp |
| Drug Approval Date (Mutual Recognition Procedure) | Calculate launch delays | Heads of Medicines Agencies | http://mri.cts-mrp.eu/Human/about |
| Incidence of Disease | Calculate market size for each market (therapeutic area and country) | Global Burden of Disease Study | http://ghdx.healthdata.org/gbd-results-tool |
| Exchange Rates | Adjust sales to Euros | European Central Bank | https://www.ecb.europa.eu/stats/policy_and_exchange_rates/euro_reference_exchange_rates/html/index.en.html |
| GDP and Population | Regression controls and market size construction | Eurostat | http://ec.europa.eu/eurostat/web/national-accounts/data/database |

An observation in our data is identified by a product, country, and year. A product is defined as a molecule, firm, and therapeutic class. Table 4 reports the main variables included in the dataset.[3] Below we describe the procedure we used to calculate each of them.

**Corp_aggr** This variable is what we use to identify the firms that are marketing a specific product. Across the EEA, more than one firm can market the same brand product. This occurs mainly because the firm that received approval sells the marketing rights for specific countries to a different company.[4] We also sometimes see multiple firms marketing the same product within the same country. Usually, these are instances of parallel traders, which we can easily identify because they first appear at least a few quarters after the original product has already been launched. However,

[3]We omit several variables for the sake of brevity. Some, like form or strength, are additional product characteristics that we lift from the original data in a straightforward way. Others, like market share, we calculate directly from the variables we do include in our list.

[4]This can occur for a variety of reasons. Sometimes, small companies may lack the necessary infrastructure to sell their drug in certain countries (this is more likely to occur in small Eastern European countries), so they license their product to a larger, established company. Other times, a company may license their product to a local company who can navigate the approval process more efficiently.

Table 4: MAIN VARIABLES IN THE DATASET USED FOR ESTIMATION

| Variable Name | Description | Variable Type |
|---|---|---|
| country | Country | Panel variable 1 |
| year | Year | Panel variable 2 |
| mlist | Active ingredient | Product definition |
| corp_aggr | List of all marketing firms | Product definition |
| lotherclass2 | Therapeutic class | Product definition |
| authorization_dt | Date of EEA-wide marketing approval | Product characteristics |
| launch_date | Date of launch in country | Product characteristics |
| end_exclusivity | Date of generic entry | Product characteristics |
| main_sample | Indicator for main sample | Product characteristics |
| main_dynamic | Indicator for dynamic sample | Product characteristics |
| sales_adjusted | Revenue sales | Sales information |
| standard_units | Volume sales | Sales information |
| mkt_size | Market size in country, year, and therapeutic class | Sales information |

we also see some rare instances of co-marketing agreements, whereby two firms sell the same product at the same time in the same country. Since we do not have information on who holds the marketing rights in a given country, we assume that the first company to launch a product is the one holding the marketing rights. The corp_aggr variable then lists all firms that hold the marketing rights in at least one country.

**Lotherclass2** This variable indicates the therapeutic class. Broadly speaking, our therapeutic classes are defined at the Anatomical Therapeutic Classification (ATC) 3 level, which corresponds to a therapeutic-pharmacological subgroup within one of the 14 main systems (Table 5). In addition, we aggregate classes that contain classes of products that share the same molecules (e.g. D7A&D7B | CORTICOSTEROID, TOPICAL, PLAIN & COMBO is a therapeutic class that combines both plain and combination corticosteroids), and classes of products that have broad applications (e.g. oncologics, which are separated in three large classes: cytotoxics, hormonals, and targeted therapies). We also separate classes of products that have similar pharmacological profile but are used to combat different diseases (e.g. J5B, which includes non-HIV antivirals, is separated in J5B1, Viral Hepatitis; J5B3, Herpes Antivirals; J5B4, Flu Antivirals; and J5B5,9, Other Respiratory Antivirals). Finally, we make a few adjustments for complex diseases with therapies that come from a broad spectrum of therapeutic areas (e.g. Multiple Sclerosis includes Beta interferons (the L3B2 ATC4 code), plus a handful of drugs from other classes, mainly monoclonal antibodies).

Table 5: ATC MAIN SYSTEMS

| Code | Contents |
|---|---|
| A | Alimentary tract and metabolism |
| B | Blood and blood forming organs |
| C | Cardiovascular system |
| D | Dermatologicals |
| G | Genito-urinary system and sex hormones |
| H | Systemic hormonal preparations, excluding sex hormones and insulins |
| J | Antiinfectives for systemic use |
| L | Antineoplastic and immunomodulating agents |
| M | Musculo-skeletal system |
| N | Nervous system |
| P | Antiparasitic products, insecticides and repellents |
| R | Respiratory system |
| S | Sensory organs |
| V | Various |

This table describes the classification of the first digit of the ATC code (or ATC1). Each letter roughly corresponds to an organ system. For more details on the ATC classification system, see http://www.ephmra.org/classification

**Authorization_dt** The authorization date variable identifies the date when a product received marketing approval for the entire EEA. We create this variable by aggregating information from a variety of sources. First, we looked on the website of the European Medicines Agency (EMA) to find drugs that received approval directly from the EMA. The marketing authorization date for these products is the date published on the EMA's website. Then, we looked on the website of the Heads of Medicines Agencies (HMA) to identify products that received a marketing authorization through the mutual recognition process. The HMA is a network of the heads of the National Competent Authorities whose organizations are responsible for the regulation of medicinal products for human and veterinary use in the European Economic Area. On its website, the HMA lists all mutual recognition applications filed for each drug. We use the date of the earliest successful mutual recognition application as the date of marketing authorization.[5] After matching our data to the EMA and HMA data, we are left with a handful of drugs that still do not have an authorization date. In these cases, we impute the authorization date as the date of the fifth launch. We believe

[5]In practice, mutual recognition applications are rarely for all European countries. However, since these applications are rarely challenged, we believe that choosing to apply only to specific countries is a strategic decision that is equivalent to the choice to delay entry. Firms may only apply for mutual recognition in a given country only when it makes sense for them to actually launch there.

this is a conservative estimate, as virtually all drugs for which we have data have authorization dates that predate the fifth launch.

**Launch_date** This variable reports the date when the product was first launched in a country. It is largely based on the variable "localproductlaunchdate" in the original IMS database. The only two imputation steps we implement are to anticipate or delay the launch date to the earliest available quarter with positive sales in the rare instances when the quarter of the launch date does not match the quarter when sales first appeared.[6]

**End_exclusivity** This variable identifies the time when a product loses exclusivity and becomes subject to generic competition. This is an important date in our model because it determines the end of the dynamic problem of the firm. To calculate this variable we first take the variable "estpatentexpirydate" from the IMS data. We then define the loss of exclusivity (LOE) date by country as the earliest of "estpatentexpirydate" and the date of first generic entry in that country. In most cases, this calculation yields a unique LOE date for a single product.[7] In all other cases, we use as the LOE date the latest of the LOE dates in France, Spain, and Italy. We do this because France, Spain, and Italy are the largest countries that use ERP. Once a product loses exclusivity in all three of these countries, we believe that the incentives to manage entry will become quite small. Consistent with this idea, we find that only a very small fraction of launches occur after our estimated date of end of exclusivity.

**Main_sample** The main sample variable identifies all products that we include in our analysis of price and demand. The sample includes prescription drugs (i.e. excluding over-the-counter products) that satisfy the following requirements:

- First launched on or after January 1st, 1995
- Patent expiration occurred on or after January 1st, 2003 (giving at least one full year of potential observations)
- At least one launch in a country on or after January 1st, 2002
- The firm selling the product is classified as the originator firm (i.e. the first recorded launch date of the molecule is associated to that firm)
- At least 50% of the launches occurred prior to patent expiration (this condition eliminates a handful of products that did not receive EU-wide approval)
- The product satisfies at least one of these three conditions:
  - It received approval from the EMA

[6] Note that we can only perform this imputation step for products whose launch date is on or after January 1st 2002.

[7] Patents generally expire at the same time in all EEA member states, so it is common for drugs to lose exclusivity in the same year in all EEA member states.

- It received at least one approval using the Mutual Recognition Procedure
- It is sold in at least 11 countries as of December 31st, 2012

481 products satisfy these requirements.

**Main_dynamic** This variable identifies a subset of main sample products whose end_exclusivity date occurs on December 2012 or earlier. We use this smaller sample (87 products in total) in the dynamic analysis.

**Sales_adjusted** and **standard_units** We calculated these variables by aggregating spending and volumes for each product on a yearly basis. In doing this aggregation we sum over sales numbers for all possible formulations of a product.

**mkt_size** This variable represents our estimate of the market size in each therapeutic area, country, and year. It is expressed in standard units (the same units used to report drug volume sales). Since we do not have direct data on market size—which includes individuals who *could* be prescribed a drug—we approximate it using data on population size and incidence rates. We describe our procedure below, in steps.

1. First, we match each product to a specific disease. We do so, by using a dataset that maps ATC 4-digit codes (Anatomical Therapeutic Chemical classification) to diseases listed on the Global Burden of Disease data. This map was kindly provided to us by the authors of Costinot et al. (2019), prior to that paper's publication. In the mapping, each product has a unique ATC-4 code, and each ATC-4 code maps to a unique disease. However, the map does not include all ATC-4 codes, so we are unable to match a small percentage of drugs (~3.5%) in three therapeutic areas: N1A | ANAESTHETICS, GENERAL, N1B | ANAESTHETICS, LOCAL, and N2A | NARCOTIC ANALGESICS.

2. Next, we map each therapeutic areas to the diseases associated with all products in that therapeutic area. This map is time- and country-invariant. In other words, we use the diseases associated to all drugs in the therapeutic area, including ones that have not been approved yet, and ones that have been approved but may not have been launched yet. Our reasoning here is that future users of not-yet-launched drugs should still be counted as potential patients.

3. We then use the "prevalence" variable from the GBD data to calculate that number of potential patients in a given therapeutic area. To do so, we simply sum the prevalence of all diseases associated with that therapeutic area. Note that the prevalence variable in the GBD data is year- and country-specific, so our estimated prevalence also varies over years and countries. In the three therapeutic areas listed above, where we cannot match disease data to drugs, we use the entire population over 15 years of age as an estimate of prevalence.

4. Finally, we need to express the prevalence variable (whose unit is number of patients) to standard units. We do this to be able to calculate market shares (our drug quantity data is expressed in standard units). The right conversion rate that we should use is the average number of standard units that a patient uses in one year of treatment. Unfortunately, there is no available estimate for this number, so we have to come up with an alternative conversion rate. To do so we calculate the ratio of total quantity sold to prevalence for each therapeutic class, country, and year. Formally, denote

$$r_{akt}^{\text{conv}} = \frac{\sum_{j \in a} q_{jkt}}{P_{akt}}$$

where $q_{jkt}$ is total standard units sold by drug $j$ in country $k$ and year $t$, and $P_{akt}$ is the prevalence of therapeutic area $a$ in country $k$ and year $t$ (i.e. the estimated number of potential customers for that class, country, and year). $r_{akt}^{\text{conv}}$ would represent the average standard units used by a patient in one year of treatment *if all patients were taking medications*. Obviously, not all potential patients will be undergoing treatment in a given year. Therefore, $r_{akt}^{\text{conv}}$ represents a lower bound for the true conversion rate. Assuming that the conversion rate is the same across countries and years, the best possible approximation we can get for the true conversion rate is the maximum of $r_{akt}^{\text{conv}}$ for each class, i.e.

$$\bar{r}_{a}^{\text{conv}} = \max_{k,t} r_{akt}^{\text{conv}}$$

Then, to get a market size estimate we calculate

$$M_{akt} = \left( \sum_{j \in a} q_{jkt} \right) \times \bar{r}_{a}^{\text{conv}} \times 1.01$$

The 1.01 multiplier ensures that the outside option has positive market share—as dictated by the underlying assumptions of the logit model we use to estimate demand.

Figure 1 plots the distribution of market share for the outside option that we recover using this procedure.

### A.4 Imputation of missing data

We need to account for three types of instances of missing data. First, some products are missing some years of data, or have negative sales recorded in a small number of instances. Second, a handful of countries have partially or completely missing data for the hospital channel. Third, the Netherlands and Sweden have missing data for a few years in the initial part of the sample.

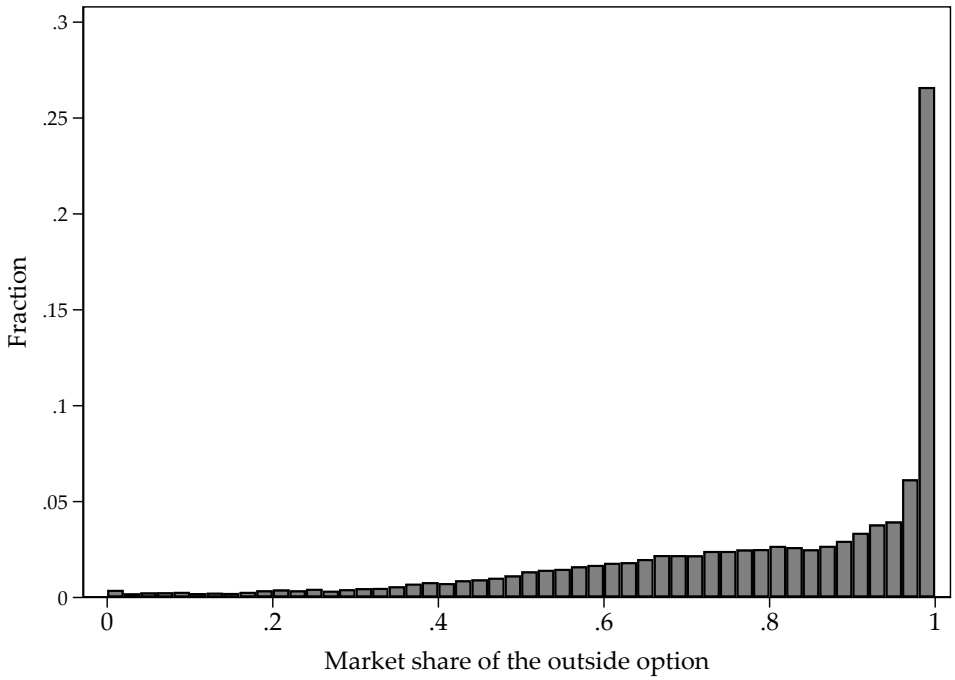



Figure 1: Distribution of the market share for the outside option

### A.4.1 Imputation of missing years of data for individual products

Occasionally, we see some years of missing data or negative sales for some drugs.[8] This can happen in three instances. First, in some cases we observe a launch date, but no sales (e.g. a product is recorded as having been launched in 2002 but first sale is recorded in 2004). Second, in other cases we may see a product disappear (e.g. it is sold every year up to 2007, but never after that). Finally, we may see sales recorded in non-consecutive years (e.g. 2003 and 2005), but not in years in between (e.g. 2004).

In the first instance, we adjust the launch date instead of the sales variable, and replace it with the date when sales are first recorded. In the second instance, we apply no adjustment and simply assume that the product exited exogenously (i.e. the firm did not choose to withdraw the product). We manually checked all these cases, and they all fall in one of two categories: either (i) the product had been declining in sales for a few years, until it disappeared; or (ii) the product was subject to a forced withdrawal by the EMA (we note these instances in Table 6).

In the last instance, we use linear interpolation to fill in the missing years. This imputation is potentially problematic. Data could be missing for two reasons: either the data were not collected for a specific drug in a particular year, or the drug did not record any sales. Our imputation strategy is appropriate in the former case, but not in the latter. We chose to impute the data this way because these occurrences are largely confined to Eastern European countries. If we mistakenly

[8] Negative sales can occur when pharmacies or hospitals reconcile inventory stock at the end of the year. We treat negative sales as missing when we do this imputation.

Table 6: WITHDRAWN PRODUCTS

| Molecule Name | Approval Date | Withdrawal Year |
|---|---|---|
| DROTRECOGIN ALFA (ACTIVATED) | 22aug2002 | 2011 |
| EPOETIN DELTA | 01jan2007 | 2009 |
| GLIMEPIRIDE#ROSIGLITAZONE | 01nov2006 | 2010 |
| LUMIRACOXIB | 01jan2005 | 2007 |
| METFORMIN#ROSIGLITAZONE | 20oct2003 | 2010 |
| RIMONABANT | 19jun2006 | 2009 |
| ROSIGLITAZONE | 11jul2000 | 2010 |
| SIBUTRAMINE | 01apr2001 | 2010 |
| SITAXENTAN | 10aug200 | 2010 |
| VALDECOXIB | 27mar2003 | 2008 |

This table lists all molecule that received approval by the EMA and later had that approval rescinded for safety reasons. An exact withdrawal date is not available since recalls generally take place over at least a few months.

attribute non-existing sales to products sold in Eastern Europe, our model will overestimate the incentive of firms to launch in Eastern Europe, which goes against the result that we find. Hence, our choice is conservative in that regard.

### A.4.2 Imputation of missing data from the hospital channel

Ireland, Lithuania, Portugal, and Romania are missing hospital sales for a few years at the beginning of the sample, but all have at least a few years of hospital sales available. Estonia, Greece, Latvia, Luxembourg, and the Netherlands are missing hospital sales entirely. We impute hospital sales using information contained in other years and other countries.

Wherever possible, we calculate drug-country specific share of sales through the hospital channels and use it to project sales for missing years. This strategy is only available for countries that have partially missing hospital sales. For countries that are missing hospital sales entirely, we use the drug-specific share of hospital sales.

The imputation works in two steps. First, we calculate the share of hospital sales $MS^h_{jk}$ (or $MS^h_j$, for the drug-specific share) using data from other years and countries. Then, hospital sales are equal to

$$\text{Sales}^{\text{Hospital}}_{jkt} = \text{Sales}^{\text{Retail}}_{jkt} \cdot \frac{MS^h_{jk}}{1 - MS^h_{jk}}$$

We do the same imputation for quantity sold and revenue sales.

For countries that are missing hospital sales in all years, we cannot impute sales of products that are sold exclusively through the hospital channel. This is a shortcoming of the data, but

unfortunately we cannot address it.

### A.4.3 Imputation of missing years of data for specific countries

The only information we have on the fully missing years for Sweden and the Netherlands is what products are sold during those years, which we can deduct using the launch date.[9] To impute quantity sold we use the demand primitives implied by our structural demand system. From our demand system we recover the parameters that determine the mean utility of each product $j$ in Sweden and the Netherlands, namely $\alpha_{j\text{SWEDEN}}$, $\alpha_{j\text{NETHERLANDS}}$, $\beta_j$, and $\eta_j$. We also recover $\sigma$ , which is the parameter that governs the covariance of the error term across the various nests. The other variables that go into calculating the predicted market share are the age of the product—which we can recover from the launch date—and the number of competitors—which we can also build by looking at which product where launched before that year. From the predicted market shares, we then calculate total sales using the market size variable.

To impute prices, we use the first available price and assume it remained constant since launch.

### A.4.4 Imputation of price data

The prices recorded in the MIDAS database are best understood as average ex-factory prices, calculated as total revenue received by the firm in a given period divided by total units sold in the same period. These prices can fluctuate for reasons that are not related to changing pricing guidelines at the government level. For example, revenues may be recorded at a different time than units sold. In other occasions, firms may provide some units of a yet unapproved drug at no charge for compassionate use.

This introduces some measurement error in the data. Since our model will work best if measurement error is limited, we run the price data through some sanity checks to parse out prices that seem unlikely to be accurate. Broadly speaking, we encounter two types of suspicious price patterns. The first one is a one-off movement in price for a given country. Usually this occurs in the first year during which a product is sold and is more likely to occur when the product is introduced late in the year. What we generally observe is a very low price, which then jumps up immediately in the following year. There are two explanations for this. One is that payments may have occurred in the following fiscal year, so they are not reflected in the current year data (but quantity is). The other is that some fraction of the product sold was used under compassionate use rules, which basically mean that the company provided it for free to certain patients even before it had officially completed the negotiation process.[10] The second one is a price for a given country that is completely inconsistent with the prices charged in the rest of Europe. This pattern is more likely to occur in Eastern Europe. Compassionate use might be the reason for these prices. When

[9] If a product that was available in a missing year disappeared before it could be recorded in our data we will not be able to observe it. However, product withdrawals are rare.

[10] For more details on compassionate use, see https://www.ema.europa.eu/en/human-regulatory/research-development/compassionate-use, retrieved March 2020.

Table 7: OBSERVATIONS FLAGGED FOR PRICE BY COUNTRY

| Country | Number of observations | Flagged |
|---|---|---|
| AUSTRIA | 3,327 | 0 |
| BELGIUM | 2,867 | 1 |
| BULGARIA | 1,841 | 6 |
| DENMARK | 3,003 | 0 |
| ESTONIA | 1,516 | 23 |
| FINLAND | 3,067 | 7 |
| FRANCE | 3,107 | 5 |
| GERMANY | 3,632 | 2 |
| GREECE | 2,487 | 8 |
| HUNGARY | 2,423 | 0 |
| IRELAND | 3,032 | 10 |
| ITALY | 3,068 | 2 |
| LATVIA | 1,799 | 10 |
| LITHUANIA | 1,909 | 28 |
| LUXEMBOURG | 2,073 | 0 |
| NETHERLANDS | 1,614 | 3 |
| NORWAY | 3,104 | 4 |
| POLAND | 2,545 | 27 |
| PORTUGAL | 2,893 | 15 |
| ROMANIA | 1,906 | 4 |
| SLOVENIA | 2,289 | 11 |
| SPAIN | 3,225 | 3 |
| SWEDEN | 2,776 | 1 |
| SWITZERLAND | 3,075 | 0 |
| UK | 3,375 | 14 |
| Total | 65,953 | 184 |

prices are that much lower in Eastern Europe, we suspect that they cannot be used in reference pricing functions.

The exact set of rules that we use to select potential measurement error in prices are the following:

1. Price values that deviate from the current average European price by a factor of 3 (either higher or lower), AND are the result of an abrupt price change (i.e. the previous year's price and the following year's price are in line with European averages)

2. Flag price values that deviate from the current European average by a factor of 10

3. Extend the flag to price values in previous or following years if the price is within 10% of the flagged value.

We then clear the flag if 4 or more countries have flagged prices for a given molecule (because this means that it is more likely that there is something real going on in the market). The remaining

flagged observations are imputed as follows. For one-off flags, we replace them with the average of the price in the previous and following year (when only one price is available, as in the case of a one-off price in the launch year, we simply use the price in the following year). Whenever all prices for a country and drug are flagged, we exclude that country and drug altogether (i.e. we assume the product was never launched there). In total, we flag 184 values out of more than 65,000. Table 7 shows the distribution of flagged values across countries.

# B Additional Background on the European Pharmaceutical Market

In this section, we present some evidence on the impact that the introduction of the EMA had on delays. We also include the reference pricing matrices that we constructed for each year.

## B.1 The Effect of the European Medicines Agency on Delays

Prior to 1995, pharmaceutical companies seeking to sell prescription drugs in Europe needed to separately apply for marketing approval in each country. The EMA was founded in 1995 to reduce the administrative burden faced by companies.

Figure 2 plots the number of new molecular entities approved each year since 1995. We distinguish between approvals obtained directly from the EMA, through the centralized procedure, approvals obtained through the Mutual Recognition Procedure (which was also installed after the EMA was introduced), and other approvals. We can see that within a few years virtually every drug is using one of the two main procedures introduced by the EMA. In recent years, drugs are also increasingly reliant on the centralized procedure, rather than the Mutual Recognition Procedure.

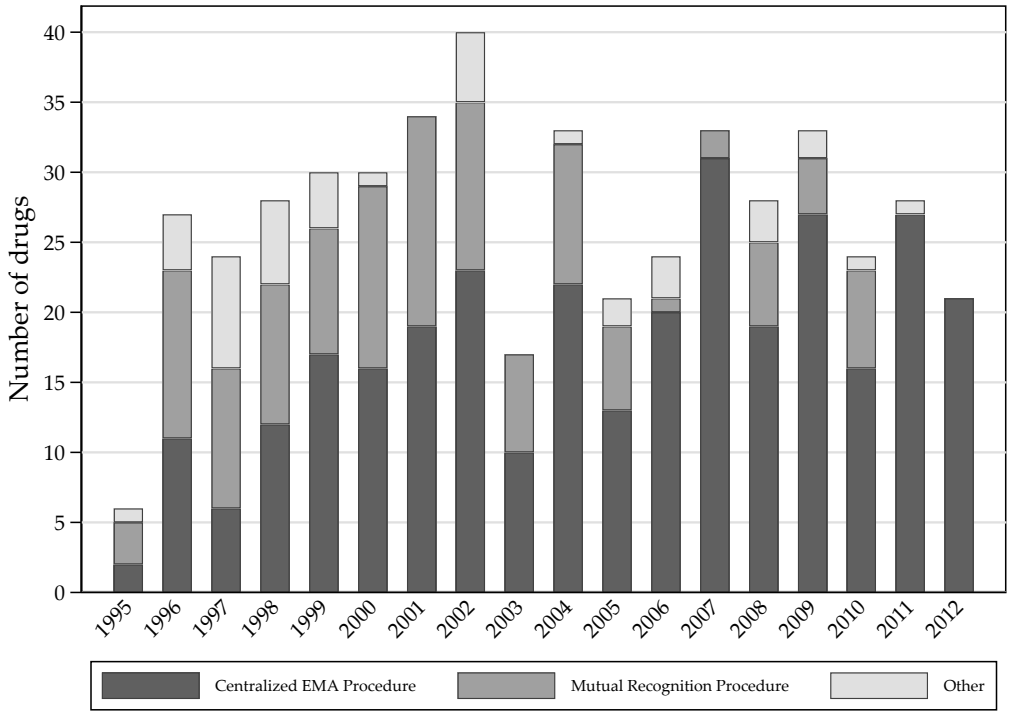
In the data, we observe a market decrease in the average launch delay before and after the founding of the EMA (Figure 3). A variety of factors likely contributes to this decrease. First, the new centralized approval procedure should reduce the fixed cost of entry in each country. This should be especially helpful for small low-income countries. Second, it should reduce the time necessary to receive approval. Third, it may reduce the protectionist tendency of member states to prevent entry from firms whose drugs compete with drugs produced by domestic manufacturers.

The clear discontinuity in delays between the pre-1995 sample, and the post-1995 sample is what motivated our choice to restrict the analysis to drugs that were first launched after January 1st, 1995.

## B.2 External Reference Pricing Functions, 2002-2012

We constructed reference pricing functions by combining several published sources (Carone et al., 2012; European Federation of Pharmaceutical Industries and Associations, 2014; Kanavos et al., 2011; Leopold et al., 2012; Wilsdon et al., 2013) with unpublished IMS reports. Even though

Figure 2: NUMBER OF DRUG APPROVALS BY YEAR

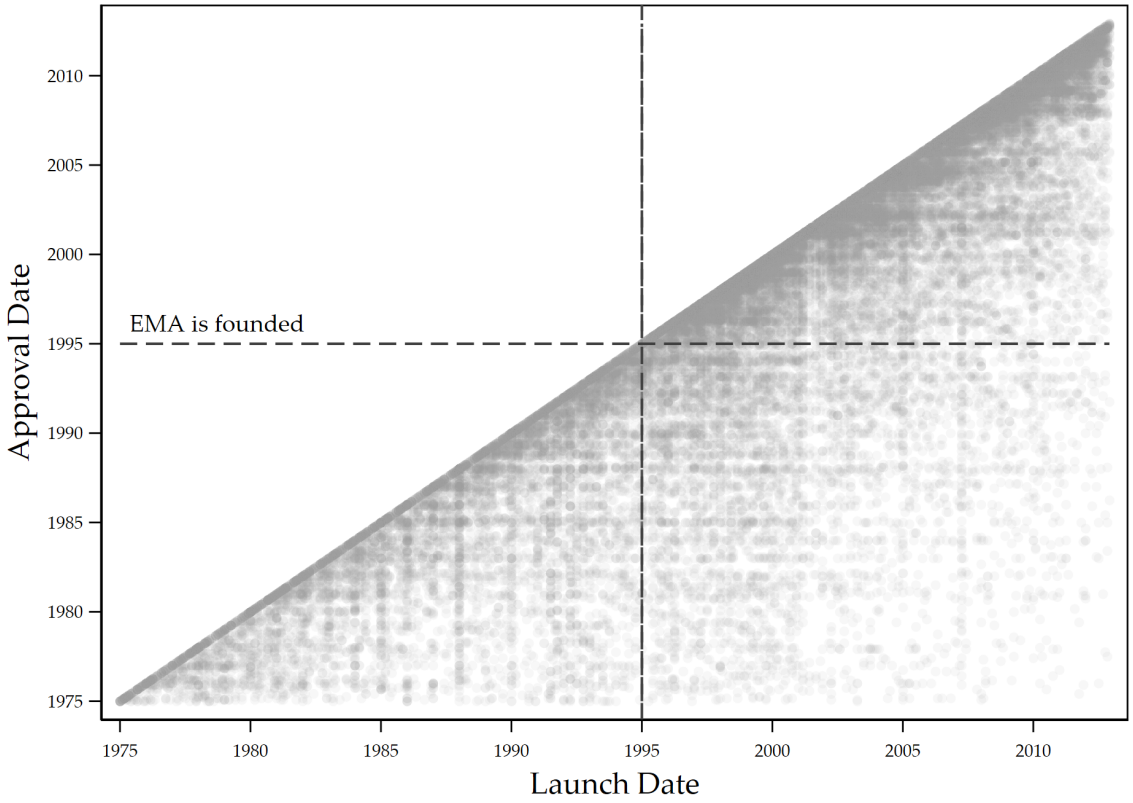



This stacked bar chart shows the number of new molecular entities approved in Europe by year, starting in 1995 (the year the EMA was founded). Approvals are divided according to the procedure used by the firm (for more details on each procedure, please consult Section 2.1 of the paper).

external reference pricing functions differ significantly across countries, only a few changes occurred to these functions over the period between 2002 and 2012, We list them in Table 8. Only four countries made changes that did not coincide with entry of new member states in the EU or Eurozone: Belgium (changed reference formula in 2011), Greece (changed basket and formula in 2010), Poland (changed basket in 2012) and Portugal (changed basket in 2007 and 2012). This reinforces our belief that most changes to the ERP function can be considered exogenous. We report reference pricing functions for all countries and years in Figure 4.

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 insurance 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). Table 9 provides an overview of these additional characteristics.

Figure 3: DELAY PATTERNS BEFORE AND AFTER THE EMA




This scatter-plot shows the distribution of the launch of new drugs. Each launch in each country is recorded separately, and is defined by two coordinates: the date in which the product was approved, and the date in which it was launched. If there were no delays, then all points should lie on the 45-degree line. It's very clear from the plot that drugs approved after 1995 experienced fewer delays (this is clear from the increased concentration of dots close to the 45-degree line), which suggests that the EMA had a large negative effect on delays. This result is what motivated our choice to focus only on the post-1995 sample.

Table 8: CHANGES TO REFERENCE PRICING FUNCTIONS

| Year | Country | Implemented change |
|---|---|---|
| 2005 | AUSTRIA | Added new EU member states (CZ EE HU LT LV PL SL SK). |
| 2008 | AUSTRIA | Added new EU member states (BG RO). |
| 2005 | BELGIUM | Added new EU member states (CZ EE HU LT LV PL SL SK). |
| 2008 | BELGIUM | Added new EU member states (BG RO). |
| 2011 | BELGIUM | Changed formula to average. |
| 2005 | FINLAND | Added new EU member states (CZ EE HU LT LV PL SL SK). |
| 2008 | FINLAND | Added new EU member states (BG RO). |
| 2010 | GREECE | Removed DK, EE, and SE. Changed formula to average of 3 lowest prices. |
| 2011 | HUNGARY | Added BG, CH, DK, EE, LT, LU, LV, NL, NO, RO, SE, UK. |
| 2005 | ITALY | Added new EU member states (CZ EE HU LT LV PL SL SK). |
| 2008 | ITALY | Added new EU member states (BG RO). |
| 2012 | POLAND | Added AT, BG, EE, FI, LV, NO, RO, SL, SK. |
| 2007 | PORTUGAL | Added Greece. |
| 2012 | PORTUGAL | Changed basket to Italy, Spain, Slovenia. |
| 2008 | SPAIN | Added Slovenia (new EURO member). |
| 2009 | SPAIN | Added Cyprus and Malta (new EURO members). |
| 2010 | SPAIN | Added Slovakia (new EURO member). |
| 2012 | SPAIN | Added Estonia, Latvia, and Lithuania (new EURO members). |

This table lists all the changes that were made to reference pricing functions, sorted by country and year. Most changes involved the addition of new EU/EURO Member States. Figure 4 also plots the full reference pricing matrices for all countries and years. Please refer to it for more details.

Table 9: ADDITIONAL ERP DETAILS FOR EEA COUNTRIES

| Country | Price referenced | ERP used for | Frequency of re-referencing (months) [a] |
|---|---|---|---|
| Austria | MNF | Reimbursed drugs | |
| Belgium | MNF | Reimbursed drugs | Undefined |
| Bulgaria | MNF | Prescription drugs | 6 |
| Switzerland | MNF | | 36 |
| Czech Republic | MNF | All drugs | 36 |
| Germany | | | |
| Denmark | | | |
| Estonia | MNF | Reimbursed drugs | 6 |
| Greece | MNF | All drugs | 6 |
| Spain | MNF | Innovative drugs | |
| Finland | PPP | Reimbursed drugs | Up to 60 |
| France | MNF | Innovative drugs | 60[b] |
| Hungary | PPP | Reimbursed drugs | 12 |
| Ireland | MNF | Innovative drugs | 24 |
| Italy | MNF | Reimbursed drugs | 24 |
| Lithuania | MNF | Reimbursed drugs | 12 |
| Luxembourg | MNF | All drugs | 18 |
| Latvia | MNF | Reimbursed drugs | 12 |
| Netherlands | PPP | Prescription drugs | 6 |
| Norway | PPP | Prescription drugs | 12 |
| Poland | MNF | Reimbursed drugs | 24 |
| Portugal | MNF | Prescription drugs | 12 |
| Romania | MNF | Reimbursed drugs | 12 |
| Sweden | | | |
| Slovenia | MNF | Reimbursed drugs | 6 |
| Slovakia | MNF | Reimbursed drugs | 6 |
| United Kingdom | | | |

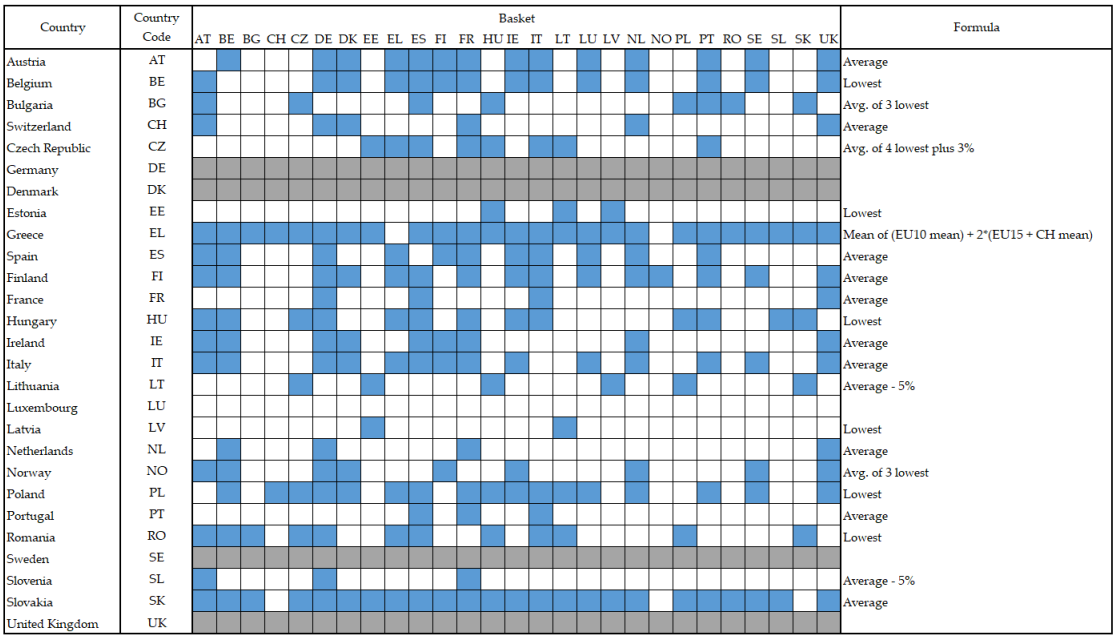
[a] European Federation of Pharmaceutical Industries and Associations. 2014. "Principles for Application of International Reference Pricing Systems." EFPIA Position Paper.

[b] France uses ERP at launch, and then guarantees the agreed upon price for five years. ERP is updated yearly after that (IMS Pharmaceutical Pricing & Reimbursement Concise Guide, France; December 2012)

We did not include these features in our model because we lacked data to do so. First, we do not know exactly when reference prices are updated, so we chose a simple rule (a one-period delay) that is roughly consistent with the average updating frequency across European countries (which is about 12 months). Second, we do not know the exact details of how PPP is applied. Finally, we do not know which drugs ERP applied to. Our parameter $\mu_k$ captures some of this cross-country variation.

## Figure 4: External Reference Pricing Matrices

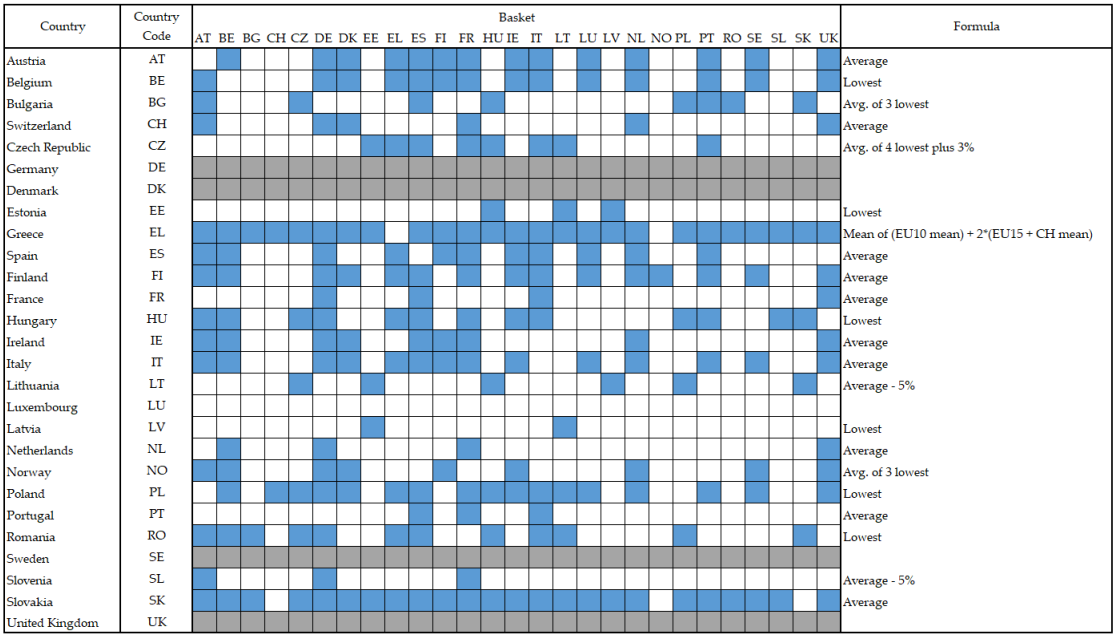
(a) 2002



| Country | Country Code | Basket | Formula |
|---|---|---|---|
| | | AT BE BG CH CZ DE DK EE EL ES FI FR HU IE IT LT LU LV NL NO PL PT RO SE SL SK UK | |
| Austria | AT | | Average |
| Belgium | BE | | Lowest |
| Bulgaria | BG | | Avg. of 3 lowest |
| Switzerland | CH | | Average |
| Czech Republic | CZ | | Avg. of 4 lowest plus 3% |
| Germany | DE | | |
| Denmark | DK | | |
| Estonia | EE | | Lowest |
| Greece | EL | | Mean of (EU10 mean) + 2*(EU15 + CH mean) |
| Spain | ES | | Average |
| Finland | FI | | Average |
| France | FR | | Average |
| Hungary | HU | | Lowest |
| Ireland | IE | | Average |
| Italy | IT | | Average |
| Lithuania | LT | | Average - 5% |
| Luxembourg | LU | | |
| Latvia | LV | | Lowest |
| Netherlands | NL | | Average |
| Norway | NO | | Avg. of 3 lowest |
| Poland | PL | | Lowest |
| Portugal | PT | | Average |
| Romania | RO | | Lowest |
| Sweden | SE | | |
| Slovenia | SL | | Average - 5% |
| Slovakia | SK | | Average |
| United Kingdom | UK | | |

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

(b) 2003

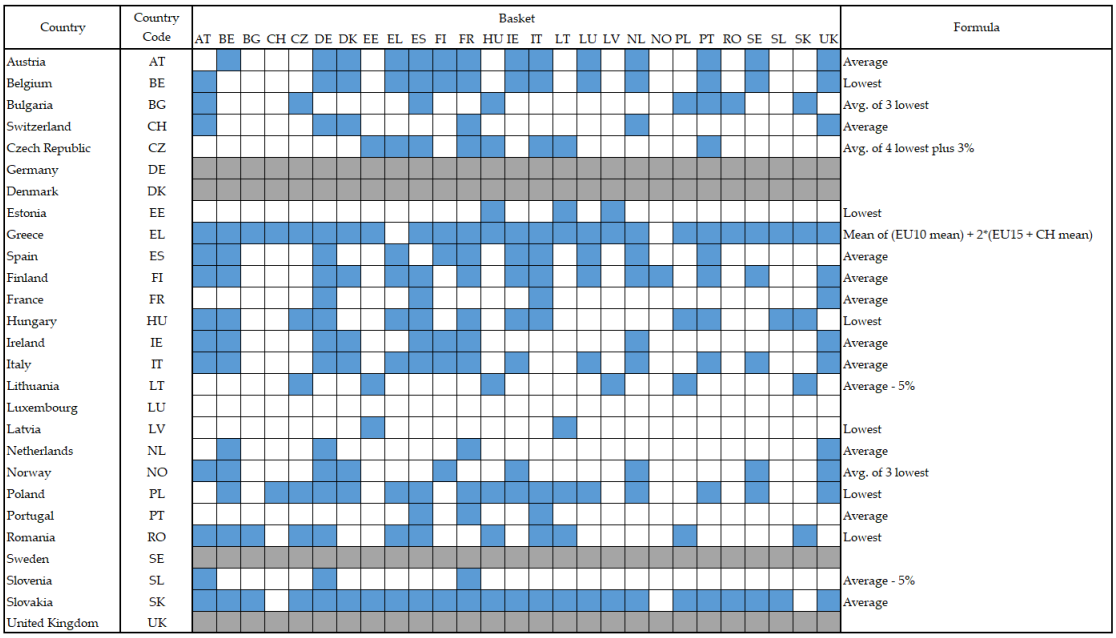


| Country | Country Code | Basket | Formula |
|---|---|---|---|
| | | AT BE BG CH CZ DE DK EE EL ES FI FR HU IE IT LT LU LV NL NO PL PT RO SE SL SK UK | |
| Austria | AT | | Average |
| Belgium | BE | | Lowest |
| Bulgaria | BG | | Avg. of 3 lowest |
| Switzerland | CH | | Average |
| Czech Republic | CZ | | Avg. of 4 lowest plus 3% |
| Germany | DE | | |
| Denmark | DK | | |
| Estonia | EE | | Lowest |
| Greece | EL | | Mean of (EU10 mean) + 2*(EU15 + CH mean) |
| Spain | ES | | Average |
| Finland | FI | | Average |
| France | FR | | Average |
| Hungary | HU | | Lowest |
| Ireland | IE | | Average |
| Italy | IT | | Average |
| Lithuania | LT | | Average - 5% |
| Luxembourg | LU | | |
| Latvia | LV | | Lowest |
| Netherlands | NL | | Average |
| Norway | NO | | Avg. of 3 lowest |
| Poland | PL | | Lowest |
| Portugal | PT | | Average |
| Romania | RO | | Lowest |
| Sweden | SE | | |
| Slovenia | SL | | Average - 5% |
| Slovakia | SK | | Average |
| United Kingdom | UK | | |

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

## Figure 4: External Reference Pricing Matrices

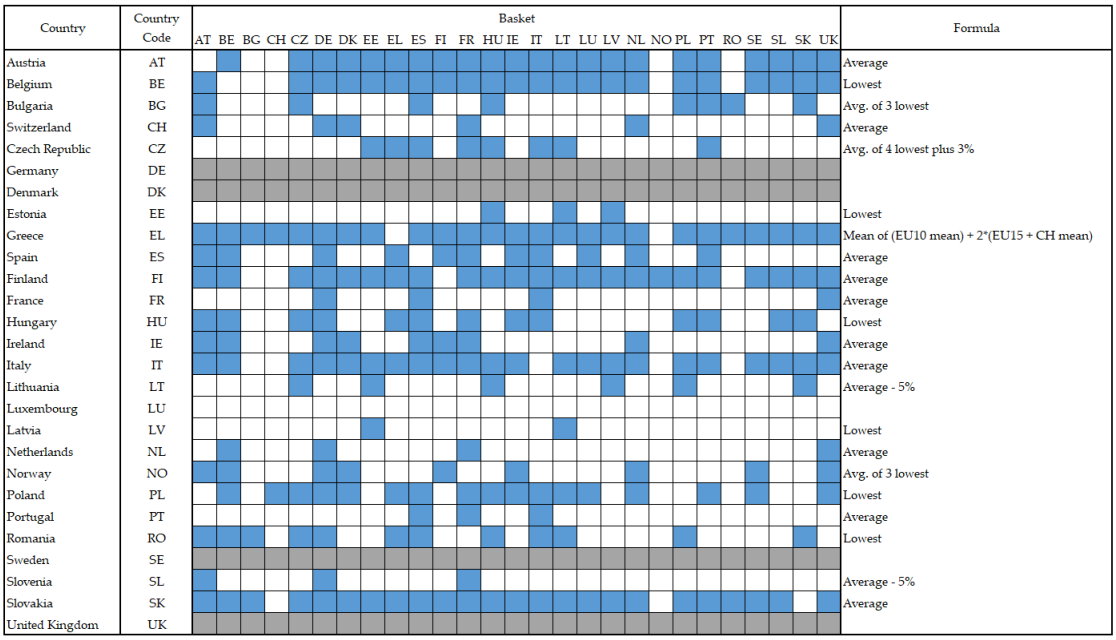
(c) 2004



| Country | Country Code | Basket | | | | | | | | | | | | | | | | | | | | | | | | | Formula |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AT | BE | BG | CH | CZ | DE | DK | EE | EL | ES | FI | FR | HU | IE | IT | LT | LU | LV | NL | NO | PL | PT | RO | SE | SL | SK | UK | |
| Austria | AT | | ■ | | | | ■ | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | | ■ | | | ■ | | ■ | | | ■ | Average |
| Belgium | BE | ■ | | | | | ■ | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | | ■ | | | ■ | | ■ | | | ■ | Lowest |
| Bulgaria | BG | ■ | | | | ■ | | | | | ■ | | | ■ | | | | | | | | ■ | ■ | ■ | | | ■ | Avg. of 3 lowest |
| Switzerland | CH | ■ | | | | | ■ | ■ | | | | | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Czech Republic | CZ | | | | | | | | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | | | | | | ■ | | | | | | Avg. of 4 lowest plus 3% |
| Germany | DE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Denmark | DK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Estonia | EE | | | | | | | | | | | | | ■ | | | ■ | | ■ | | | | | | | | | | Lowest |
| Greece | EL | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Mean of (EU10 mean) + 2*(EU15 + CH mean) |
| Spain | ES | ■ | ■ | | | | ■ | | ■ | | | ■ | ■ | | ■ | ■ | | ■ | | ■ | | | ■ | | | | | ■ | Average |
| Finland | FI | ■ | ■ | | | | ■ | ■ | | ■ | ■ | | ■ | | ■ | ■ | | ■ | | ■ | ■ | | ■ | | ■ | | | ■ | Average |
| France | FR | | | | | | ■ | | | ■ | | | | | | ■ | | | | | | | | | | | | ■ | Average |
| Hungary | HU | ■ | ■ | | | ■ | ■ | | | ■ | ■ | | ■ | | ■ | ■ | | | | | | ■ | ■ | | | ■ | ■ | | Lowest |
| Ireland | IE | ■ | ■ | | | | ■ | ■ | | | ■ | ■ | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Italy | IT | ■ | ■ | | | | ■ | ■ | | ■ | ■ | ■ | ■ | | ■ | | | ■ | | ■ | | | ■ | | ■ | | | ■ | Average |
| Lithuania | LT | | | | | ■ | | | ■ | | | | | ■ | | | | | ■ | | | ■ | | | | | ■ | | Average - 5% |
| Luxembourg | LU | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Latvia | LV | | | | | | | | ■ | | | | | | | | ■ | | | | | | | | | | | | Lowest |
| Netherlands | NL | | ■ | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | ■ | Average |
| Norway | NO | ■ | ■ | | | | ■ | ■ | | | | ■ | | | ■ | | | | | ■ | | | | | ■ | | | ■ | Avg. of 3 lowest |
| Poland | PL | | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | | | ■ | | ■ | | | ■ | Lowest |
| Portugal | PT | | | | | | | | | | ■ | | ■ | | | ■ | | | | | | | | | | | | | Average |
| Romania | RO | ■ | ■ | ■ | | ■ | ■ | | | ■ | ■ | | | ■ | | ■ | ■ | | | | | | ■ | | | | | ■ | Lowest |
| Sweden | SE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slovenia | SL | ■ | | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | | Average - 5% |
| Slovakia | SK | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | | ■ | Average |
| United Kingdom | UK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

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

(d) 2005

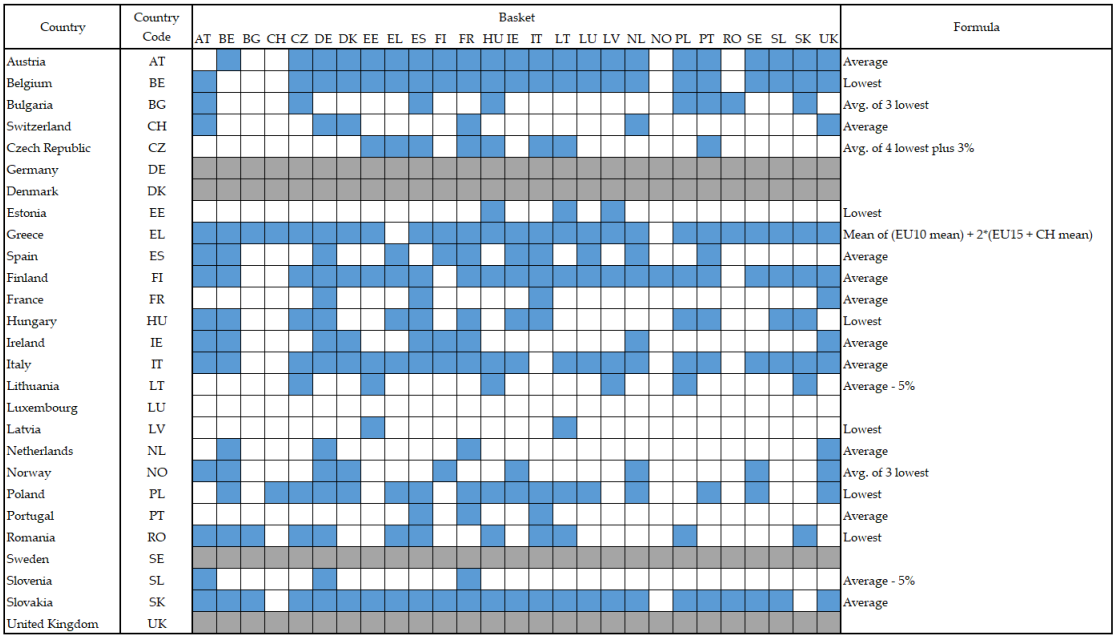


| Country | Country Code | Basket | | | | | | | | | | | | | | | | | | | | | | | | | Formula |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AT | BE | BG | CH | CZ | DE | DK | EE | EL | ES | FI | FR | HU | IE | IT | LT | LU | LV | NL | NO | PL | PT | RO | SE | SL | SK | UK | |
| Austria | AT | | ■ | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | Average |
| Belgium | BE | ■ | | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | Lowest |
| Bulgaria | BG | ■ | | | | ■ | | | | | ■ | | | ■ | | | | | | | | ■ | ■ | ■ | | | ■ | ■ | Avg. of 3 lowest |
| Switzerland | CH | ■ | | | | | ■ | ■ | | | | | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Czech Republic | CZ | | | | | | | | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | | | | | | ■ | | | | | | Avg. of 4 lowest plus 3% |
| Germany | DE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Denmark | DK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Estonia | EE | | | | | | | | | | | | | ■ | | | ■ | | ■ | | | | | | | | | | Lowest |
| Greece | EL | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Mean of (EU10 mean) + 2*(EU15 + CH mean) |
| Spain | ES | ■ | ■ | | | | ■ | | ■ | | | ■ | ■ | | ■ | ■ | | ■ | | ■ | | | ■ | | | | | ■ | Average |
| Finland | FI | ■ | ■ | | | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | Average |
| France | FR | | | | | | ■ | | | ■ | | | | | | ■ | | | | | | | | | | | | ■ | Average |
| Hungary | HU | ■ | ■ | | | ■ | ■ | | | ■ | ■ | | ■ | | ■ | ■ | | | | | | ■ | ■ | | | ■ | ■ | | Lowest |
| Ireland | IE | ■ | ■ | | | | ■ | ■ | | | ■ | ■ | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Italy | IT | ■ | ■ | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | Average |
| Lithuania | LT | | | | | ■ | | | ■ | | | | | ■ | | | | | ■ | | | ■ | | | | | ■ | | Average - 5% |
| Luxembourg | LU | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Latvia | LV | | | | | | | | ■ | | | | | | | | ■ | | | | | | | | | | | | Lowest |
| Netherlands | NL | | ■ | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | ■ | Average |
| Norway | NO | ■ | ■ | | | | ■ | ■ | | | | ■ | | | ■ | | | | | ■ | | | | | ■ | | | ■ | Avg. of 3 lowest |
| Poland | PL | | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | | | ■ | | ■ | | | ■ | Lowest |
| Portugal | PT | | | | | | | | | | ■ | | ■ | | | ■ | | | | | | | | | | | | | Average |
| Romania | RO | ■ | ■ | ■ | | ■ | ■ | | | ■ | ■ | | | ■ | | ■ | ■ | | | | | | ■ | | | | | ■ | Lowest |
| Sweden | SE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slovenia | SL | ■ | | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | | Average - 5% |
| Slovakia | SK | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | | ■ | Average |
| United Kingdom | UK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

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

## Figure 4: External Reference Pricing Matrices

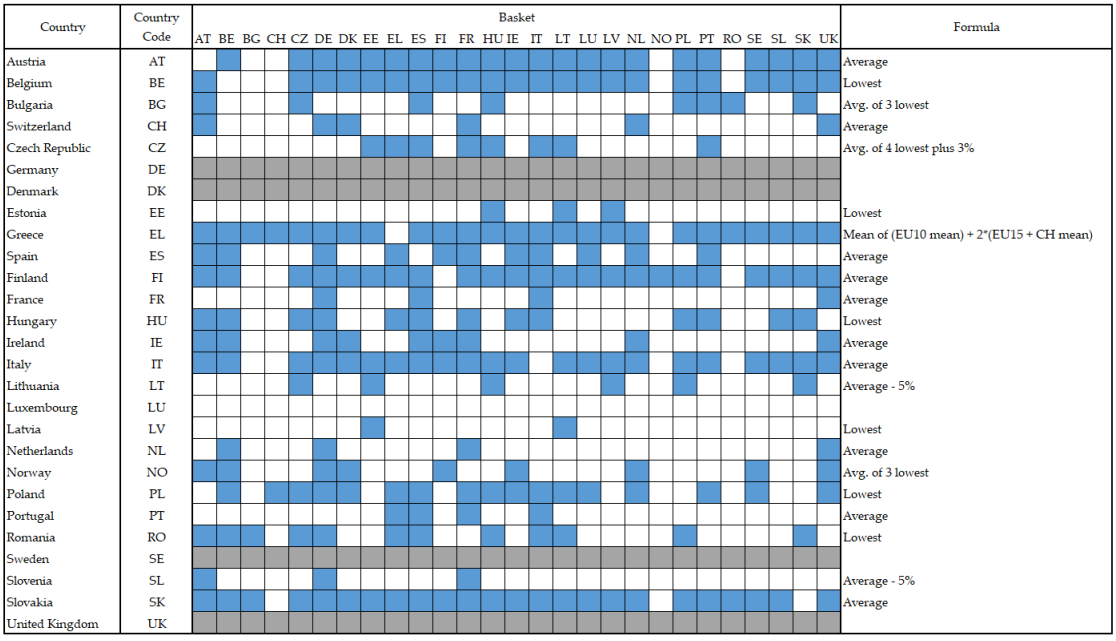
(e) 2006



| Country | Country Code | Basket | | | | | | | | | | | | | | | | | | | | | | | | | Formula |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AT | BE | BG | CH | CZ | DE | DK | EE | EL | ES | FI | FR | HU | IE | IT | LT | LU | LV | NL | NO | PL | PT | RO | SE | SL | SK | UK | |
| Austria | AT | | ■ | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | Average |
| Belgium | BE | ■ | | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | Lowest |
| Bulgaria | BG | ■ | | | | ■ | | | | | ■ | | | ■ | | | | | | | | ■ | ■ | ■ | | | ■ | | Avg. of 3 lowest |
| Switzerland | CH | ■ | | | | | ■ | ■ | | | | | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Czech Republic | CZ | | | | | | | | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | | | | | | ■ | | | | | | Avg. of 4 lowest plus 3% |
| Germany | DE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Denmark | DK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Estonia | EE | | | | | | | | | | | | | ■ | | | ■ | | ■ | | | | | | | | | | Lowest |
| Greece | EL | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Mean of (EU10 mean) + 2*(EU15 + CH mean) |
| Spain | ES | ■ | ■ | | | | ■ | | | ■ | | | ■ | | ■ | ■ | | ■ | | ■ | | | ■ | | | | | | Average |
| Finland | FI | ■ | ■ | | | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | Average |
| France | FR | | | | | | ■ | | | | ■ | | | | | ■ | | | | | | | | | | | | ■ | Average |
| Hungary | HU | ■ | ■ | | | ■ | ■ | | | ■ | | ■ | ■ | | ■ | ■ | | | | | | ■ | ■ | | | ■ | ■ | | Lowest |
| Ireland | IE | ■ | ■ | | | | ■ | ■ | | | ■ | ■ | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Italy | IT | ■ | ■ | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | Average |
| Lithuania | LT | | | | | ■ | | | ■ | | | | | ■ | | | | | ■ | | | ■ | | | | | ■ | | Average - 5% |
| Luxembourg | LU | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Latvia | LV | | | | | | | | ■ | | | | | | | | ■ | | | | | | | | | | | | Lowest |
| Netherlands | NL | | ■ | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | ■ | Average |
| Norway | NO | ■ | ■ | | | | ■ | ■ | | | | ■ | | | ■ | | | | | ■ | | | | | | ■ | | | Avg. of 3 lowest |
| Poland | PL | | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | | | ■ | | ■ | | | ■ | Lowest |
| Portugal | PT | | | | | | | | | | ■ | | ■ | | | ■ | | | | | | | | | | | | | Average |
| Romania | RO | ■ | ■ | ■ | | ■ | ■ | | | ■ | ■ | | | ■ | | ■ | ■ | | | | | | ■ | | | | | ■ | Lowest |
| Sweden | SE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slovenia | SL | ■ | | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | | Average - 5% |
| Slovakia | SK | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | | ■ | Average |
| United Kingdom | UK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

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

(f) 2007

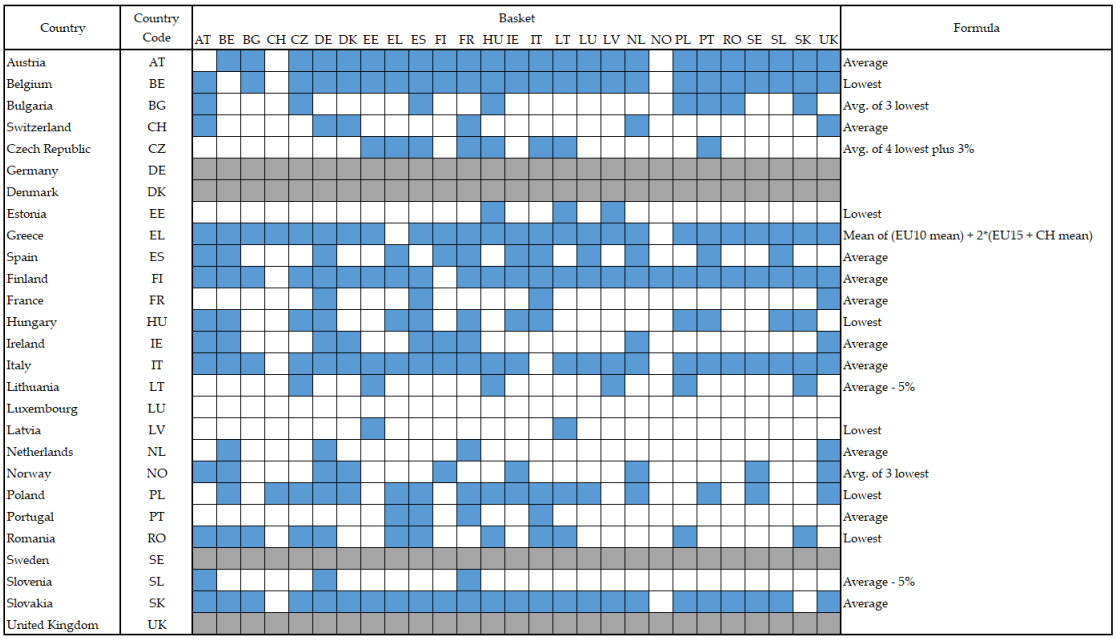


| Country | Country Code | Basket | | | | | | | | | | | | | | | | | | | | | | | | | Formula |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AT | BE | BG | CH | CZ | DE | DK | EE | EL | ES | FI | FR | HU | IE | IT | LT | LU | LV | NL | NO | PL | PT | RO | SE | SL | SK | UK | |
| Austria | AT | | ■ | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | Average |
| Belgium | BE | ■ | | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | Lowest |
| Bulgaria | BG | ■ | | | | ■ | | | | | ■ | | | ■ | | | | | | | | ■ | ■ | ■ | | | ■ | | Avg. of 3 lowest |
| Switzerland | CH | ■ | | | | | ■ | ■ | | | | | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Czech Republic | CZ | | | | | | | | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | | | | | | ■ | | | | | | Avg. of 4 lowest plus 3% |
| Germany | DE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Denmark | DK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Estonia | EE | | | | | | | | | | | | | ■ | | | ■ | | ■ | | | | | | | | | | Lowest |
| Greece | EL | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Mean of (EU10 mean) + 2*(EU15 + CH mean) |
| Spain | ES | ■ | ■ | | | | ■ | | | ■ | | | ■ | | ■ | ■ | | ■ | | ■ | | | ■ | | | | | | Average |
| Finland | FI | ■ | ■ | | | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | Average |
| France | FR | | | | | | ■ | | | | ■ | | | | | ■ | | | | | | | | | | | | ■ | Average |
| Hungary | HU | ■ | ■ | | | ■ | ■ | | | ■ | | ■ | ■ | | ■ | ■ | | | | | | ■ | ■ | | | ■ | ■ | | Lowest |
| Ireland | IE | ■ | ■ | | | | ■ | ■ | | | ■ | ■ | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Italy | IT | ■ | ■ | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | Average |
| Lithuania | LT | | | | | ■ | | | ■ | | | | | ■ | | | | | ■ | | | ■ | | | | | ■ | | Average - 5% |
| Luxembourg | LU | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Latvia | LV | | | | | | | | ■ | | | | | | | | ■ | | | | | | | | | | | | Lowest |
| Netherlands | NL | | ■ | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | ■ | Average |
| Norway | NO | ■ | ■ | | | | ■ | ■ | | | | ■ | | | ■ | | | | | ■ | | | | | | ■ | | | Avg. of 3 lowest |
| Poland | PL | | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | | | ■ | | ■ | | | ■ | Lowest |
| Portugal | PT | | | | | | | | | | ■ | | ■ | | | ■ | | | | | | | | | | | | | Average |
| Romania | RO | ■ | ■ | ■ | | ■ | ■ | | | ■ | ■ | | | ■ | | ■ | ■ | | | | | | ■ | | | | | ■ | Lowest |
| Sweden | SE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slovenia | SL | ■ | | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | | Average - 5% |
| Slovakia | SK | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | | ■ | Average |
| United Kingdom | UK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

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

## Figure 4: External Reference Pricing Matrices

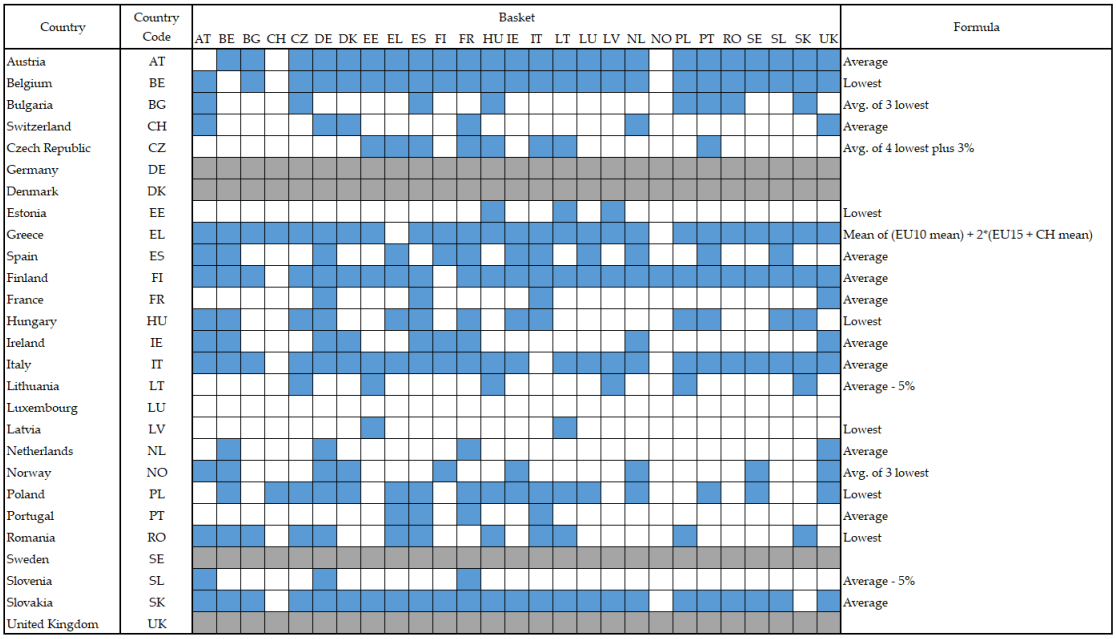
(g) 2008



| Country | Country Code | Basket | | | | | | | | | | | | | | | | | | | | | | | | | Formula |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AT | BE | BG | CH | CZ | DE | DK | EE | EL | ES | FI | FR | HU | IE | IT | LT | LU | LV | NL | NO | PL | PT | RO | SE | SL | SK | UK | |
| Austria | AT | | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Average |
| Belgium | BE | ■ | | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Lowest |
| Bulgaria | BG | ■ | | | | ■ | | | | | ■ | | | ■ | | | | | | | | ■ | ■ | ■ | | | ■ | ■ | Avg. of 3 lowest |
| Switzerland | CH | ■ | | | | | ■ | ■ | | | | | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Czech Republic | CZ | | | | | | | | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | | | | | | | ■ | | | | | Avg. of 4 lowest plus 3% |
| Germany | DE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Denmark | DK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Estonia | EE | | | | | | | | | | | | | ■ | | | ■ | | ■ | | | | | | | | | | Lowest |
| Greece | EL | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Mean of (EU10 mean) + 2*(EU15 + CH mean) |
| Spain | ES | ■ | ■ | | | ■ | | | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | | | Average |
| Finland | FI | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Average |
| France | FR | | | | | | ■ | | | | ■ | | | | | ■ | | | | | | | | | | | | ■ | Average |
| Hungary | HU | ■ | ■ | | | ■ | ■ | | | ■ | | ■ | ■ | | ■ | ■ | | | | | | ■ | ■ | | | ■ | ■ | | Lowest |
| Ireland | IE | ■ | ■ | | | | ■ | ■ | | | | ■ | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Italy | IT | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Average |
| Lithuania | LT | | | | | ■ | | | ■ | | | | | ■ | | | | | ■ | | | ■ | | | | | ■ | | Average - 5% |
| Luxembourg | LU | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Latvia | LV | | | | | | | | ■ | | | | | | | | ■ | | | | | | | | | | | | Lowest |
| Netherlands | NL | | ■ | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | ■ | Average |
| Norway | NO | ■ | ■ | | | | ■ | ■ | | | | ■ | | | ■ | | | | | ■ | | | | | | ■ | | ■ | Avg. of 3 lowest |
| Poland | PL | | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | | | ■ | | ■ | | | ■ | Lowest |
| Portugal | PT | | | | | | | | | ■ | ■ | | ■ | | | ■ | | | | | | | | | | | | | Average |
| Romania | RO | ■ | ■ | ■ | | ■ | ■ | | | ■ | ■ | | | ■ | | ■ | ■ | | | | | | ■ | | | | ■ | | Lowest |
| Sweden | SE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slovenia | SL | ■ | | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | | Average - 5% |
| Slovakia | SK | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | | ■ | Average |
| United Kingdom | UK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

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

(h) 2009

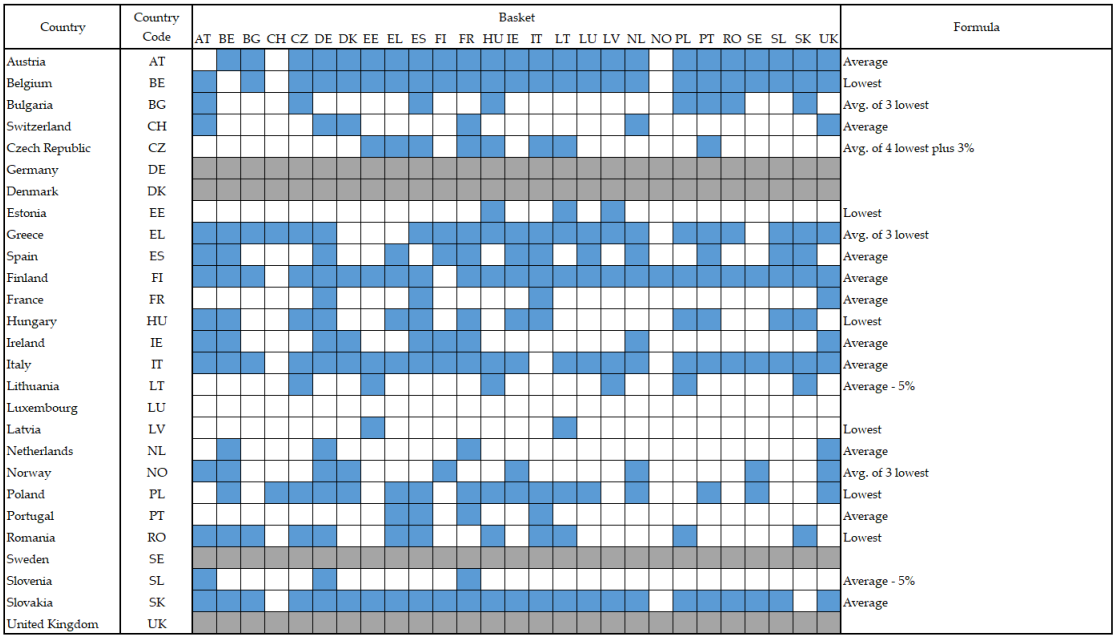


| Country | Country Code | Basket | | | | | | | | | | | | | | | | | | | | | | | | | Formula |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AT | BE | BG | CH | CZ | DE | DK | EE | EL | ES | FI | FR | HU | IE | IT | LT | LU | LV | NL | NO | PL | PT | RO | SE | SL | SK | UK | |
| Austria | AT | | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Average |
| Belgium | BE | ■ | | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Lowest |
| Bulgaria | BG | ■ | | | | ■ | | | | | ■ | | | ■ | | | | | | | | ■ | ■ | ■ | | | ■ | ■ | Avg. of 3 lowest |
| Switzerland | CH | ■ | | | | | ■ | ■ | | | | | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Czech Republic | CZ | | | | | | | | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | | | | | | | ■ | | | | | Avg. of 4 lowest plus 3% |
| Germany | DE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Denmark | DK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Estonia | EE | | | | | | | | | | | | | ■ | | | ■ | | ■ | | | | | | | | | | Lowest |
| Greece | EL | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Mean of (EU10 mean) + 2*(EU15 + CH mean) |
| Spain | ES | ■ | ■ | | | ■ | | | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | | | Average |
| Finland | FI | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Average |
| France | FR | | | | | | ■ | | | | ■ | | | | | ■ | | | | | | | | | | | | ■ | Average |
| Hungary | HU | ■ | ■ | | | ■ | ■ | | | ■ | | ■ | ■ | | ■ | ■ | | | | | | ■ | ■ | | | ■ | ■ | | Lowest |
| Ireland | IE | ■ | ■ | | | | ■ | ■ | | | ■ | ■ | ■ | | | | | | | ■ | | | | | | | | ■ | Average |
| Italy | IT | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | Average |
| Lithuania | LT | | | | | ■ | | | ■ | | | | | ■ | | | | | ■ | | | ■ | | | | | ■ | | Average - 5% |
| Luxembourg | LU | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Latvia | LV | | | | | | | | ■ | | | | | | | | ■ | | | | | | | | | | | | Lowest |
| Netherlands | NL | | ■ | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | ■ | Average |
| Norway | NO | ■ | ■ | | | | ■ | ■ | | | | ■ | | | ■ | | | | | ■ | | | | | | ■ | | ■ | Avg. of 3 lowest |
| Poland | PL | | ■ | | ■ | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | | | ■ | | ■ | | | ■ | Lowest |
| Portugal | PT | | | | | | | | | ■ | ■ | | ■ | | | ■ | | | | | | | | | | | | | Average |
| Romania | RO | ■ | ■ | ■ | | ■ | ■ | | | ■ | ■ | | | ■ | | ■ | ■ | | | | | | ■ | | | | ■ | | Lowest |
| Sweden | SE | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slovenia | SL | ■ | | | | | ■ | | | | | | ■ | | | | | | | | | | | | | | | | Average - 5% |
| Slovakia | SK | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | | ■ | Average |
| United Kingdom | UK | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

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

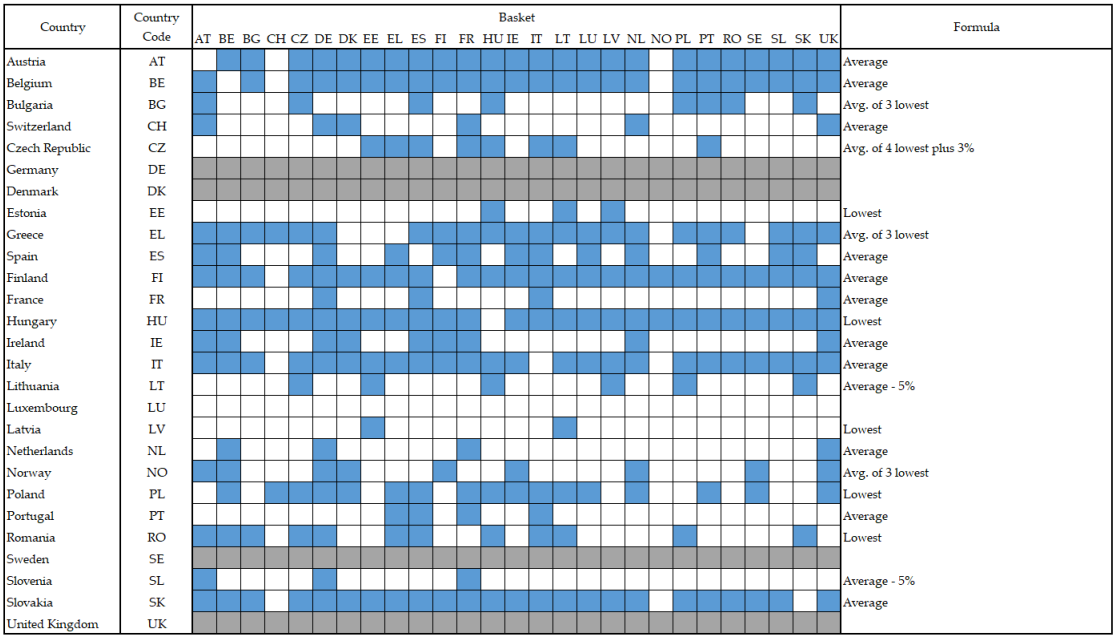
## Figure 4: External Reference Pricing Matrices

(i) 2010




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

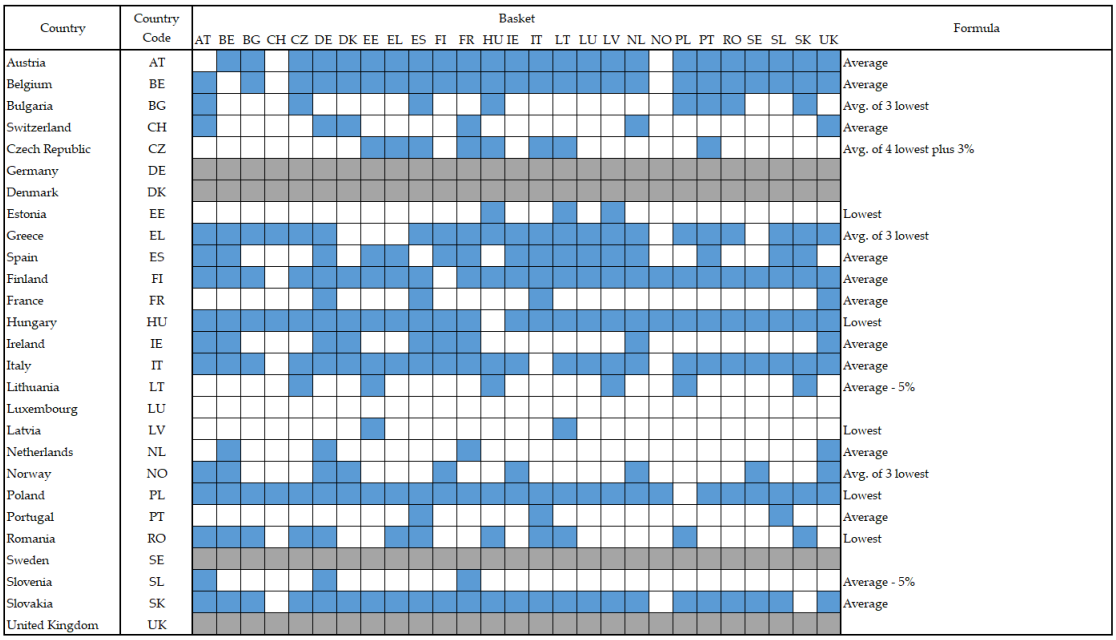
(j) 2011




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

Figure 4: EXTERNAL REFERENCE PRICING MATRICES

(k) 2012




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

# C Additional results and robustness checks

In this section, we report additional results and robustness tests for some of the model specifications. The section is divided in three subsections. In the first subsection, we explore alternative explanations for launch delays that do not involve ERP, and find that none of these alternative explanations fits all patterns in the data.

In the second subsection, we report some reduced-form evidence on the correlation between price levels and delays. One of the other possible avenues that firms can exploit to push back against reference price is increase prices in lower income countries. We find some suggestive evidence that backs this idea: price levels are strongly correlated with GDP per capita among higher-income Western European countries, but the relationship flattens significantly among Eastern European countries suggesting that firms apply some sort of price floor for European countries, possibly as a response to ERP.

In the third subsection, we report results of a placebo test of the pricing model where we assigned a fake reference function to the four countries that do not use reference pricing. The test shows that the $\mu_k$ coefficient—which measures the weight given to the reference function in the pricing equation— is zero for all four countries.

## C.1 Testing alternative explanations of launch delays

The presence of delays does not represent, in and of itself, enough evidence that firms are responding strategically to reference pricing. Delays 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. In this section, we check whether reduced form patterns in the data are consistent with these alternative explanations, and find that none of these possible alternatives can fully explain all patterns in the data.

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. 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 in country $k$. 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 $k$ as $q_k$ and $p_k$ respectively and assume WLOG that $p_1 > p_2$.

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

**External reference pricing.** First, assume that differential pricing can only be sustained for one period. After that, if the drug is available abroad, 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. 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

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

The LHS of this equation represents the revenue loss caused by reference pricing in the second period, while the RHS represents the additional sales from anticipating entry in country 2. Since the loss depends on the difference between $p_1$ and $p_2$, delays in country 2 should be inversely correlated with prices even after controlling for revenue.

**Fixed costs of entry.** Second, assume that in order to enter in a country the firm must pay a stochastic fixed cost of entry $\zeta_{kt} \sim F_{k\zeta}(\theta_{k\zeta})$. We can treat the entry problem in each country

separately. In period 1, the firm decides whether to delay or not based on

$$\max \{2p_k q_k - \zeta_{k1}; p_k q_k - \mathbb{E}[\zeta_{k2}]\}$$

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

$$\zeta_{k1} > p_k q_k - \mathbb{E}[\zeta_{k2}] \tag{2}$$

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.** Third, 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 predicts that delays should be inversely correlated with price even after accounting for revenue.

**Limited number of applications.** Finally, 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 $k$ if and only if

$$p_k q_k = \max_{\ell \in \{1,2\}} \{p_\ell q_\ell\} \tag{3}$$

As in the model with fixed costs of entry, prices shouldn't matter after accounting for revenue.

While delays can arise in all four variations of the model, each scenario predicts different delay patterns. To test which models fits the data best we run three tests. First, we check whether delays are inversely correlated with prices after controlling for revenue. As a test of this hypothesis we regress delays on revenue and price:

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

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 reference price. 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 noisier. Our results (Table 10) 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. Hence, this test rules out both fixed costs of entry and limits to the number of applications as the explanations behind delays.

Table 10: IMPACT OF PRICE AND REVENUE ON DELAYS

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| $\ln\left(\text{Yearly Rev}_{ij}\right)$ | −2.752 | −3.725 | −2.756 | −3.769 | −2.756 | −3.757 |
| | (0.092) | (0.096) | (0.092) | (0.097) | (0.092) | (0.096) |
| $\ln\left(\text{avg}\left(p_{jkt}\right)\right)$ | −0.207 | −8.782 | | | | |
| | (0.087) | (0.767) | | | | |
| $\ln\left(P_{ijt_0}\right)$ | | | −0.170 | −5.907 | | |
| | | | (0.088) | (0.726) | | |
| $\ln\left(\max\left\{p_{jkt}\right\}\right)$ | | | | | −0.163 | −8.701 |
| | | | | | (0.087) | (0.718) |
| Drug FE | Y | Y | Y | Y | Y | Y |
| Country FE | N | Y | N | Y | N | Y |
| $R^2$ | 0.09 | 0.38 | 0.09 | 0.38 | 0.09 | 0.39 |
| $N$ | 8,819 | 8,819 | 8,819 | 8,819 | 8,819 | 8,819 |

Second, we test whether capacity constraints are likely to exist. 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 approximately 22% of launches are missing at the end of the year in which peak output is achieved (Figure 5). Hence, this test suggests that capacity constraints cannot entirely explain delays.

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

$$\ln\left(p_{jkt}\right) = \theta_j + \gamma_{ka} + \delta_t + \varepsilon_{jkt}$$

where $\gamma_{ka}$ is a fixed effect for country and drug age, measured in years starting from the approval year. We also include drug fixed effects $\theta_k$ and year fixed effects $\delta_t$. 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 $\gamma_{ka}$ coefficients for Germany and those of the the three other countries (Figure 6). 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.

Figure 5: FRACTION OF LAUNCHES OCCURRING AFTER PEAK CAPACITY IS REACHED

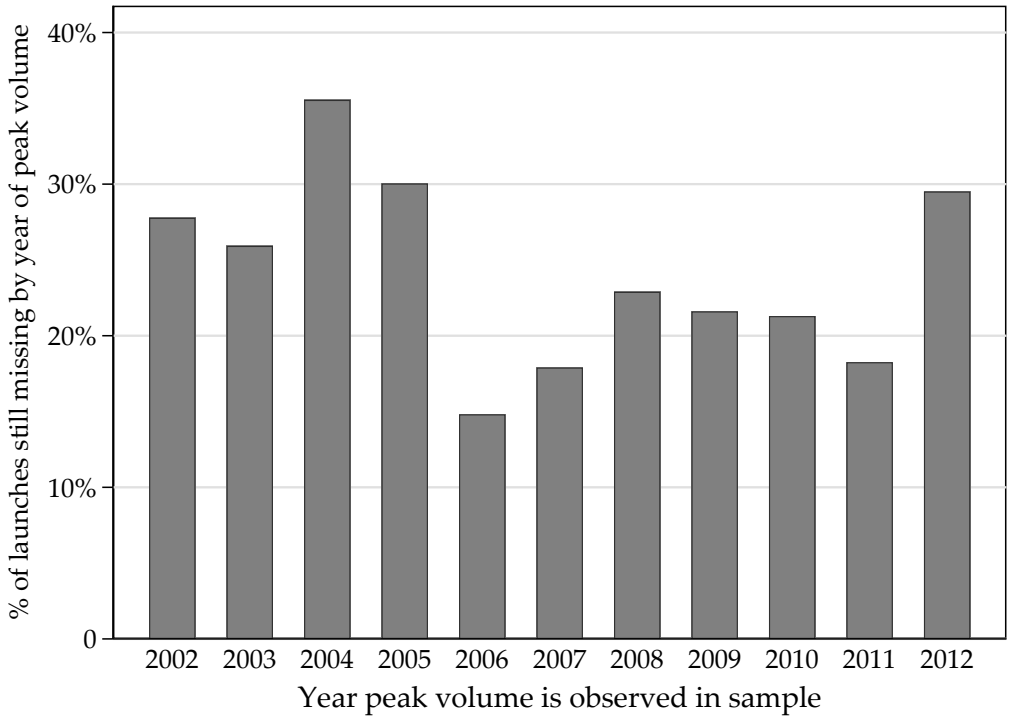
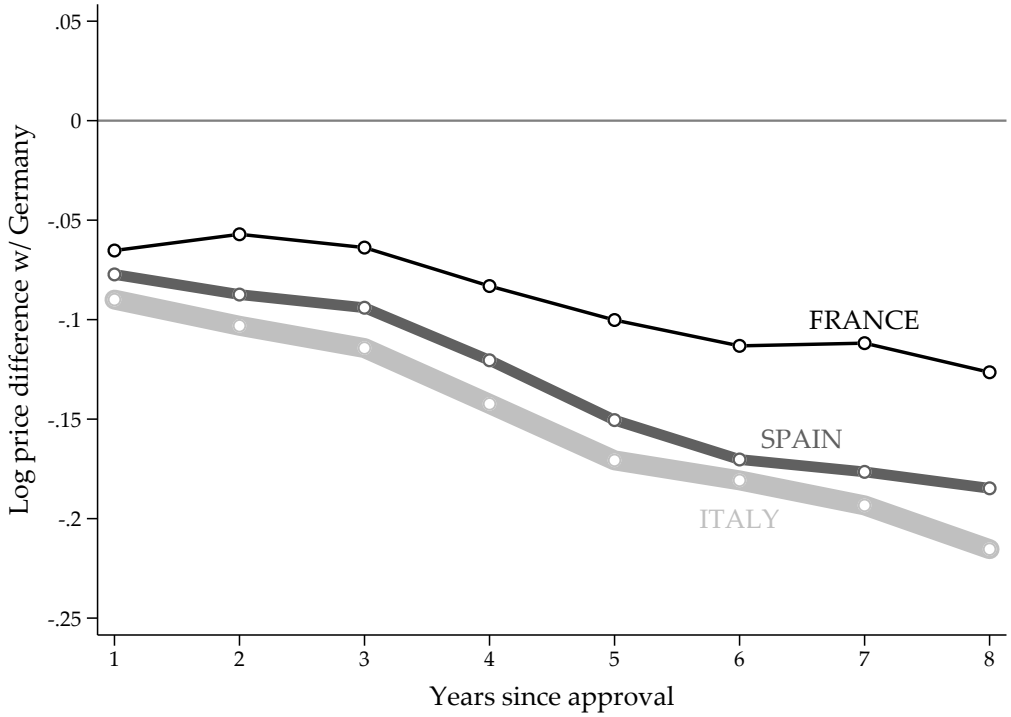



Figure 6: CHANGE IN PRICES OVER TIME AND ACROSS COUNTRIES



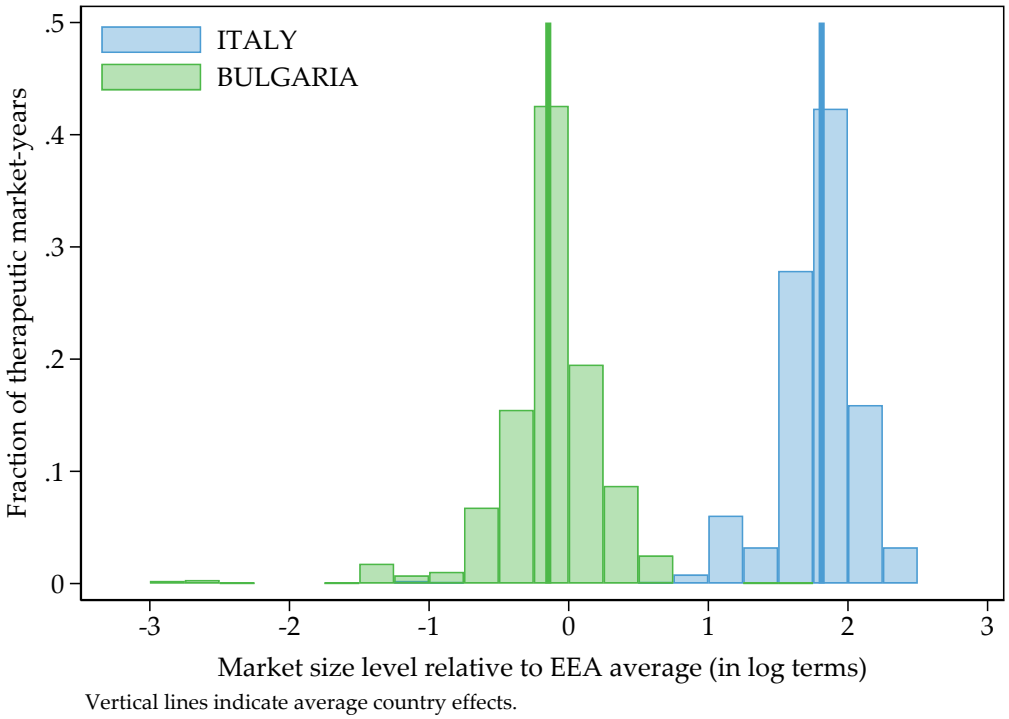



Figure 7: Market size variation within country: Italy and Bulgaria

## C.2 Additional output from the demand system

In this section, we provide additional output from the estimation of the demand system. We highlight three findings.

First, we find that there is broad heterogeneity in both market size and drug preferences within country. In other words, overall sales relative to the European average across therapeutic areas in a given country fluctuate quite a bit. Figure 7 shows variation in market size relative to the European average for Italy and Bulgaria. Each observation in the graph is a therapeutic area and year. Bulgaria's market size is roughly around the European average, while Italy's is much larger. However, there are therapeutic areas where Bulgaria's market size is well above average, and, conversely, there are therapeutic areas where Italy's market size is only slightly above average.

Moreover, even within therapeutic areas, it is often the case that specific products are more popular in certain countries than in others. Figure 8 plots the distribution of relative market share of products in the same two countries. Each observation is a drug in our main sample of 481. Once again, we see large dispersion. Bulgaria's average market share is lower, which reflects the propensity of lower-income countries to rely more heavily on generic and drugs produced for the local market only (neither of which appear in our main sample). However, the distribution shows large variance around the mean.

These two patterns are crucial for our findings: market size and drug preferences are likely known by firms in advance, and will be a key input in their pricing decisions. The ability to

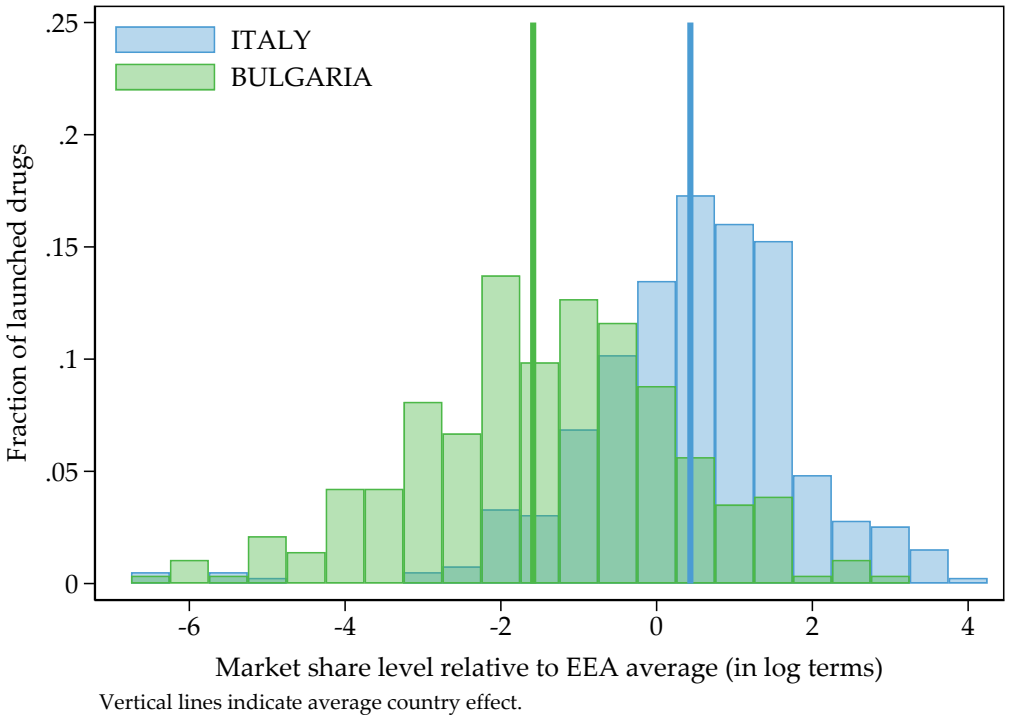



Figure 8: Drug preference variation within country: Italy and Bulgaria

estimate these in a flexible way is an important feature of our model, and motivated our choice to use drug-country specific fixed effects (as opposed to separate drug and country fixed effects).

Second, we plot the distribution of the coefficient $\beta_j$, which in the model estimates the change in demand with age (Figure 9). On average, age has a positive impact on demand, though there are a few exceptions. The distribution in the overall sample (which includes drugs with more recent authorization dates) has a longer right tail, which may reflect faster growth in demand in the first few years after launch.

Third, and final, we plot the distribution of the nesting parameters $\sigma$ (Figure 10). The nesting parameter falls outside the unit interval in around half of the therapeutic classes, though most coefficients fall in the interval between 0 and 2. This is somewhat surprising, since drugs with sharing the same molecule are probably close substitutes. However, in our model, the only type of same-molecule competition that can occur is between originator products and parallel traded products (in the estimation, we aggregate all other products at the molecule level to avoid excessive entry and exit). Parallel-traded products only exist in a handful of countries, and can be perceived as different or lower quality.

### C.3 Correlation between prices and GDP per capita

Delaying launches is only one of many possible strategic reactions to ERP. Another possible strategy that firms can use is to try to increase the prices they charge in lower-income countries. While

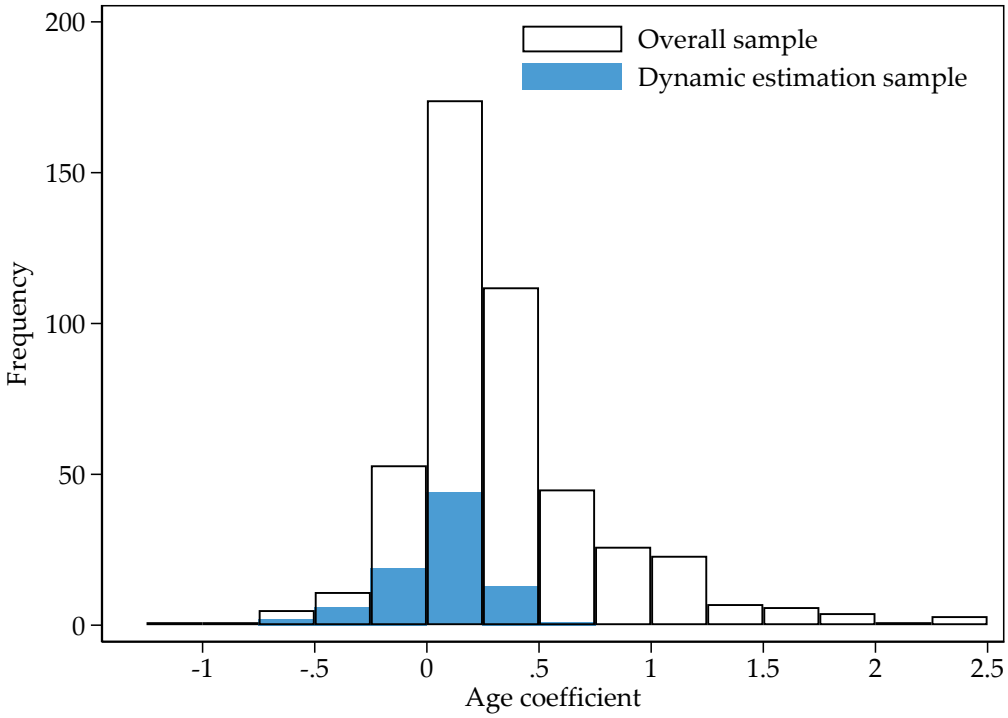



Figure 9: Impact of age on demand across drugs

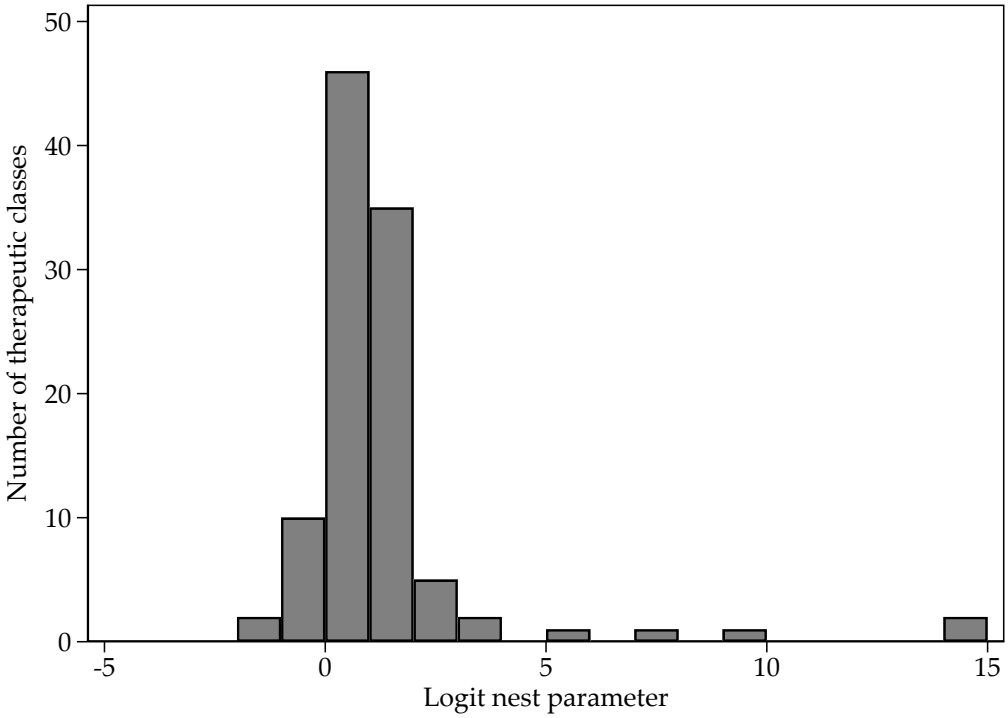



Figure 10: Distribution of the logit nest parameter

it is hard to estimate what prices in lower-income countries could be in the absence of ERP, we do find some evidence that this is the case. In Figure 11 we plot GDP per capita in 2012 against average country-level drug prices.[11] We find a strong positive correlation between the two among Western European countries. However, the relationship is significantly flatter among Eastern European countries. Not only that, but the price level in Easter Europe is on average only marginally lower than the price level in the Western European countries with the lowest drug prices (e.g. Italy, Spain). These small differences are in spite of large differences in GDP per capita. Put together, these results suggest that firms do try to charge higher-than-expected prices in lower-income European countries.

Notice however, that this does not affect the estimation or the counterfactual of our main paper. We estimate country-level price fixed effects, which capture price levels in the current equilibrium. This would be problematic if we were running counterfactual that required computing prices with a different pricing rule. Instead, our counterfactual exercise is only concerned with computing delays, which are unaffected by the change in pricing rule. As long as ERP is replaced by another policy, the optimal strategy, under relatively weak conditions, is to apply for entry in all countries at the same time.

The calculation of the lump sum transfer is equally unaffected, because it is done under the assumption that government keep the same pricing rules as in the current equilibrium.

### C.4 Correlation between quantity, prices, and delays

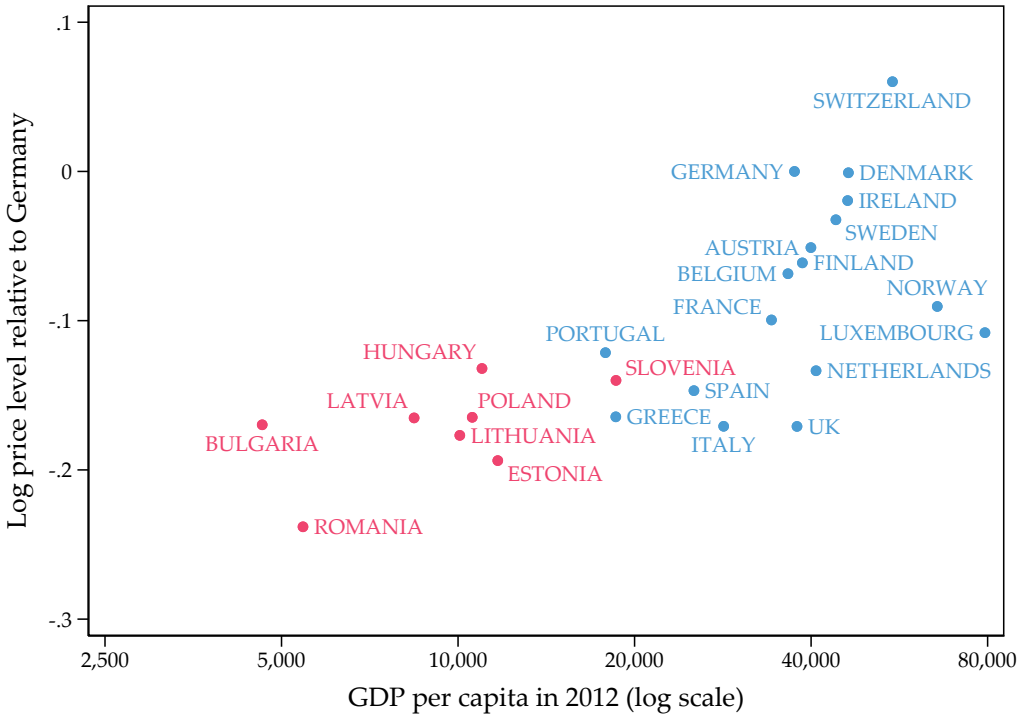
To decide what control variables to include in our models of demand and price, we tested whether within-country variation in demand and drug prices is also correlated with launch behavior. To do this, we ran regressions of volume sales and price on drug and country fixed-effects:

$$\begin{aligned} \ln\left(q_{jkt}\right) &= \theta_j^q + \gamma_k^q + \varepsilon_{jkt}^q \\ \ln\left(p_{jkt}\right) &= \theta_j^p + \gamma_k^p + \varepsilon_{jkt}^p \end{aligned}$$

In these regressions $j$ indexes products, $k$ indexes country, and $t$ indexes periods. The residuals $\varepsilon_{jkt}^q$ and $\varepsilon_{jkt}^p$ reflect both a fully stochastic component (reflecting uncertainty that the firm cannot observe in advance), as well as information that is not reflected in the drug and country fixed effect, but that the firm may see in advance. For example, if Italy had a disproportionately large number of diabetes patients, we would expect $\varepsilon_{iITAt}^q$ to be large for insulin and other diabetes medication. To check whether within-country variation in market outcomes is correlated with

[11] The country-level drug prices are coefficients from a regression of log price on drug and country fixed effects.

Figure 11: CORRELATION BETWEEN DRUG PRICES AND GDP PER CAPITA




firm behavior, we regress delays on the residuals $\varepsilon^q_{jkt^0}$ and $\varepsilon^p_{jkt^0}$ (where $t^0$ is the launch year)[12]

$$\begin{aligned} y_{jk} &= \gamma^y_k + \beta_q \varepsilon^q_{jkt^0} + \varepsilon^y_{jk} \\ y_{jk} &= \gamma^y_k + \beta_p \varepsilon^p_{jkt^0} + \varepsilon^y_{jk} \end{aligned}$$

We test both delay in months, and the order of launch (i.e. the rank of the country $k$ in the launch sequence of drug $i$).

We report our results in Table 11. We find that $\varepsilon^q_{ijt^0}$ is negatively correlated with both months of delay as well as order of launch, meaning that within each country, products with low demand tend to enter later.

Surprisingly, we do not find the same result for $\varepsilon^p_{ijt^0}$. There are two possible reasons to explain this phenomenon. First, it could be that $\varepsilon^p_{ijt^0}$ only contains information that the firm does not observe in advance. In this case, we would not expect the firm's actions to reflect anything that is unknown about price prior to entry. Second, it is possible that there is not enough within-country, cross-drug variation in price. In the data, we note that a price regression on drug and country fixed effects achieves a coefficient of variation of approximately 0.97, meaning that these two variables soak up a lot of variation. The equivalent regression on sales has a much lower coefficient (around

[12]When we do not observe sales in the launch year (because launch occurred before 2002), we use the residual from the first available year.

Table 11: CORRELATION OF WITHIN-COUNTRY VARIATION AND FIRM BEHAVIOR

| | **Launch Order** | | **Delay** | |
|---|---|---|---|---|
| $\varepsilon^q_{ijt}$ | $-0.396$ | | $-1.544$ | |
| | $[-0.443, -0.348]$ | | $[-1.758, -1.331]$ | |
| $\varepsilon^p_{ijt}$ | | $-0.089$ | | $-0.794$ |
| | | $[-0.369, 0.191]$ | | $[-2.041, 0.453]$ |
| Country FE | Y | Y | Y | Y |
| $N$ | 8872 | 8872 | 8872 | 8872 |
| $R^2$ | 0.50 | 0.48 | 0.24 | 0.22 |

0.77), which suggests that there may be more variation in quantity that the firm can act upon.

These regressions motivate our choice of demand and price functions. In our demand functions, we include drug-country fixed effects, since our regressions suggest that using a separate set of fixed effects may fail to capture important variation that is known to the firm. In our price functions however, we include drug and country fixed effects separately. The regression results in this section suggest this is enough to capture the relevant variation in price.

### C.5 Placebo tests on the pricing model

To test whether our pricing model could be simply capturing correlations between price across Europe that are generated by other confounding factors, we run a falsification test where we assign fictitious reference pricing functions to the four countries that do not use ERP: Denmark, Germany, Sweden, and the UK. If our model is correctly specified, the value of $\mu$ for these countries should be close to zero, since the reference pricing should not affect their prices.

To Denmark and Sweden, we assign the same reference function as Norway. To Germany, we assign the same reference function as France, and to the UK, we assign the same reference function as Ireland.[13]

[13]Since countries cannot reference their own price, we replace it with the price of the country whose reference function we use (e.g. since France references Germany, we replace Germany with France in the reference function used by Germany).

Table 12: FALSIFICATION TEST RESULTS

| **Country** | $\ln\left(\gamma_j\right)$ | $\mu_k$ |
|---|---|---|
| Austria | -0.103 | 0.340 |
| Belgium | -0.121 | 0.222 |
| Bulgaria | -0.200 | 1.000 |
| Denmark[b] | -0.082 | $2 \times 10^{-6}$ |
| Estonia | -0.188 | 1.000 |
| Finland | -0.132 | 0.307 |
| France | -0.094 | $2 \times 10^{-6}$ |
| Germany[a,b] | 0.000 | $2 \times 10^{-7}$ |
| Greece | -0.093 | 1.000 |
| Hungary | -0.240 | 0.998 |
| Ireland | -0.076 | 0.592 |
| Italy | -0.174 | 1.000 |
| Latvia | -0.242 | 0.902 |
| Lithuania | -0.242 | 1.000 |
| Luxembourg[c] | -0.236 | 0 |
| Netherlands | -0.206 | $1 \times 10^{-5}$ |
| Norway | -0.167 | 1.000 |
| Poland | -0.064 | 0.932 |
| Portugal | -0.194 | 1.000 |
| Romania | -0.280 | 1.000 |
| Slovenia | -0.258 | 0.769 |
| Spain | -0.158 | 1.000 |
| Sweden[b] | -0.107 | $2 \times 10^{-6}$ |
| Switzerland | -0.003 | $1 \times 10^{-6}$ |
| UK[b] | -0.192 | $5 \times 10^{-18}$ |

[a] The price level is normalized to Germany's.
[b] Denmark, Germany, Sweden, and the UK do not actually use ERP during 2002-2012.
[c] Luxembourg references the price of the country of origin of the drug. Since we do not know country of origin, we assume that $\mu_j$ equals zero.

We find that for all four countries, the $\mu$ coefficient is almost exactly 0 (Table 12). We note that this does not conclusively prove that our model does not suffer from misspecification, or other estimation issues. However, the test gives us confidence that the $\mu$ parameters are picking up the impact of ERP correctly.

### C.6 Evidence in support of the pricing model results and specification

In this section, we present a few pieces of evidence in support of the specification we chose for the price function. We present three results. First, we show reduced-form evidence that prices fall following entry in Eastern European countries, but only in countries using ERP, and more so in countries that reference Eastern Europe directly. Second, we present evidence from a few examples

Table 13: Impact of entry in Eastern European countries on prices in Italy, Spain, and Germany

| | Italy | | Spain | | Germany | |
|---|---|---|---|---|---|---|
| | **Individual regression** | **Joint regressions** | **Individual regression** | **Joint regressions** | **Individual regression** | **Joint regressions** |
| **Bulgaria** | -0.052 | -0.032 | -0.001 | 0.017 | 0.045 | 0.015 |
| | (0.013) | (0.013) | (0.013) | (0.015) | (0.011) | (0.013) |
| **Estonia** | -0.071 | -0.037 | -0.050 | -0.038 | 0.070 | 0.038 |
| | (0.013) | (0.014) | (0.015) | (0.017) | (0.013) | (0.016) |
| **Hungary** | -0.028 | 0.021 | -0.021 | 0.000 | 0.046 | 0.016 |
| | (0.020) | (0.021) | (0.017) | (0.021) | (0.012) | (0.015) |
| **Latvia** | -0.070 | -0.042 | -0.045 | -0.039 | 0.050 | 0.002 |
| | (0.013) | (0.015) | (0.017) | (0.019) | (0.013) | (0.016) |
| **Lithuania** | -0.048 | -0.003 | -0.015 | 0.009 | 0.065 | 0.036 |
| | (0.014) | (0.015) | (0.016) | (0.018) | (0.012) | (0.014) |
| **Poland** | -0.056 | -0.003 | -0.037 | -0.025 | 0.033 | 0.000 |
| | (0.015) | (0.018) | (0.017) | (0.019) | (0.011) | (0.013) |
| **Romania** | -0.028 | 0.006 | -0.006 | 0.009 | 0.040 | 0.011 |
| | (0.011) | (0.013) | (0.014) | (0.015) | (0.010) | (0.013) |
| **Slovenia** | -0.059 | -0.030 | -0.018 | -0.001 | 0.026 | -0.024 |
| | (0.017) | (0.018) | (0.019) | (0.019) | (0.012) | (0.016) |

of specific price paths, showing that prices do not seem to react very strongly to reference prices that are above them. Third, we show that the data does not fit the predictions of a more standard Bertrand Nash model.

### C.6.1 Evidence of prices falling after entry in Eastern European countries

We present some evidence that prices fall following entry in Eastern European countries. We do this by running a simple regression of log price on drug fixed effects, plus an indicator variable for entry in each Eastern European country (we also run separate regressions for each indicator, since they tend to be somewhat collinear). We test the impact on prices in three countries: Italy and Spain (both of whom appear to follow ERP closely according to our structural estimation results), and Germany (which has a similar market size and also uses the Euro, minimizing measurement issues from exchange rate fluctuation).

Table 13 reports our results. Italy shows the clearest results, which makes sense, since Italy references Eastern Europe directly. Spain also sees a negative impact, though somewhat less pronounced. This also makes sense, as Spain does not reference Eastern European countries directly. However, Spain references Greece, and Italy, both of whom include Eastern Europe in their reference function. We do not find any evidence of a negative impact on price in Germany (in fact, almost all coefficients are positive, which may even suggest a compensating effect).

In unreported results, we also tested the movement of prices following entry in Western Eu-

ropean countries, and tried to use referencing functions to isolate the impact of ERP. We did not find strong results, which we attribute to two things. First, entry in Western European countries is usually quick, and only minimally staggered. Most drugs will enter all of Western Europe within 3-4 years, meaning that it is very common to see several launches in Western Europe in the same year. This makes it hard to isolate the impact of entry in a single country. Second, according to our structural estimation, many Western European countries do not actually appear to follow their reference function very closely, so this result is in line with what our structural model suggests.

### C.6.2 Evidence of kinks in the pricing function

In the next subsection, we show examples of price paths of specific drugs, some of which show concretely how prices in different countries react differently to changes in the reference price. We provide this evidence in partial support of the choice to use a kink in the price function, and to illustrate the type of variation that helps pin down the $\mu$ parameter in the structural estimation.

Omalizumab (Xolair) was approved by the EMA on October 25, 2005. Figure 12 plots the price and reference price paths of this product in four countries, which help us highlight a few features of the relationship between price and reference price across countries.

We note three things. First, the launch sequence of this product provides a relatively clean natural experiment. Even though the EEA-wide approval was given in 2005, the product had been available in Norway since late 2004 (this is not a very common occurrence, but it does happen from time to time). Omalizumab was launched in six countries in 2005: Denmark, Germany, Netherlands, Poland, Sweden, and the UK. Hence, the first year in which a reference price was available to Norway was 2006. In that year, we observe the price falling from about €450 per unit, to €370 per unit, a drop of approximately 25% that leaves the price at very similar level to the reference price. The Norwegian price then remains at a similar level to the reference price, despite diverging slightly in the last few years. This is the kind of pattern in the data that the structural model would interpret as a sign that a country is adhering closely to its reference pricing policy (Norway has a $\mu$ coefficient of 1 in our estimation, suggesting close adherence).

Second, we can see fairly well in the price patterns of Italy and Austria that prices do not seem to react to changes in the reference price when the reference price lies above the price itself. In the case of Italy, the reference price is almost always above the price, and the price is basically constant: it grows slightly in 2007 and 2008—but then falls back to the exact level of the initial price—and it dips in 2012, which is when the reference price dips below the launch price for the first time. In Austria, the pattern is even clearer. The price is constant until 2010, which is the first year in which the reference price dips below the launch price. The price does not fully adjust to the reference price, suggesting that Austria may only use reference price as one of many inputs in price setting (accordingly, our model estimates a coefficient $\mu$ of about one third for Austria). These types of patterns, which repeat more or less consistently across many of the drugs we see, suggest that a kinked price function is more appropriate.

Third and final, we can compare the behavior of price in Austria (and, to some extent, Italy) to

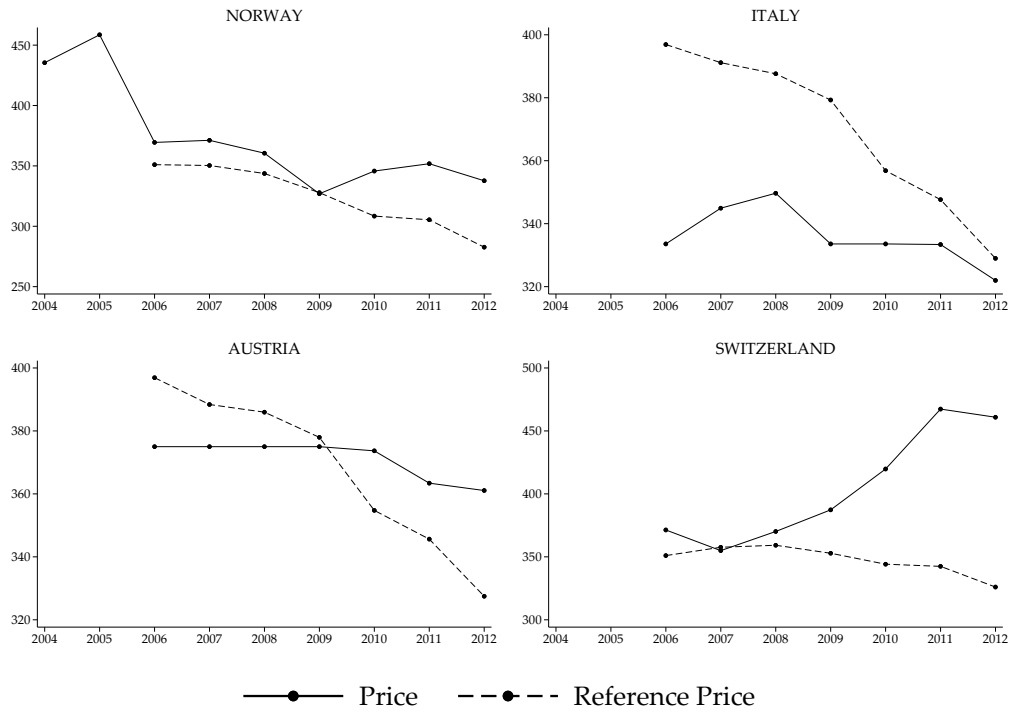



Figure 12: Price paths for Omalizumab (Xolair)

the behavior of price in Switzerland. Here, even though the reference price is below the observed price for virtually the entire sample period, the two lines are diverging. This type of behavior suggests that Switzerland does not rely on ERP, despite its stated policy. This is picked up by our model, which estimates the Swiss coefficient for $\mu$ as very close to zero.

### C.6.3 Test of Bertrand Nash predictions

Generally speaking, The use of bargaining models is widespread in the literature on prescription drugs, and the use of more traditional pricing models, such as Bertrand Nash, is rare (see for example Dubois et al. (2018) use Bertrand Nash in the US, but a bargaining model for Canada, whose regulation is much closer to that of European countries). Even in the US, where firms are subject to fewer pricing constraints and government regulation, drug pricing is often modeled as bilateral bargaining between manufacturers and PBMs, as in e.g. Feng (2020) and Berndt and Newhouse (2010).

As additional support for our modeling choice, we present here a simple test of a basic prediction of the Bertrand Nash model: prices should fall when a new competitor enters. We run a simple regression of the form:

$$\ln\left(p_{jkt}\right) = \alpha_j + \gamma_k + \theta_t + \beta N_{jkt} + \varepsilon_{jkt} \tag{4}$$

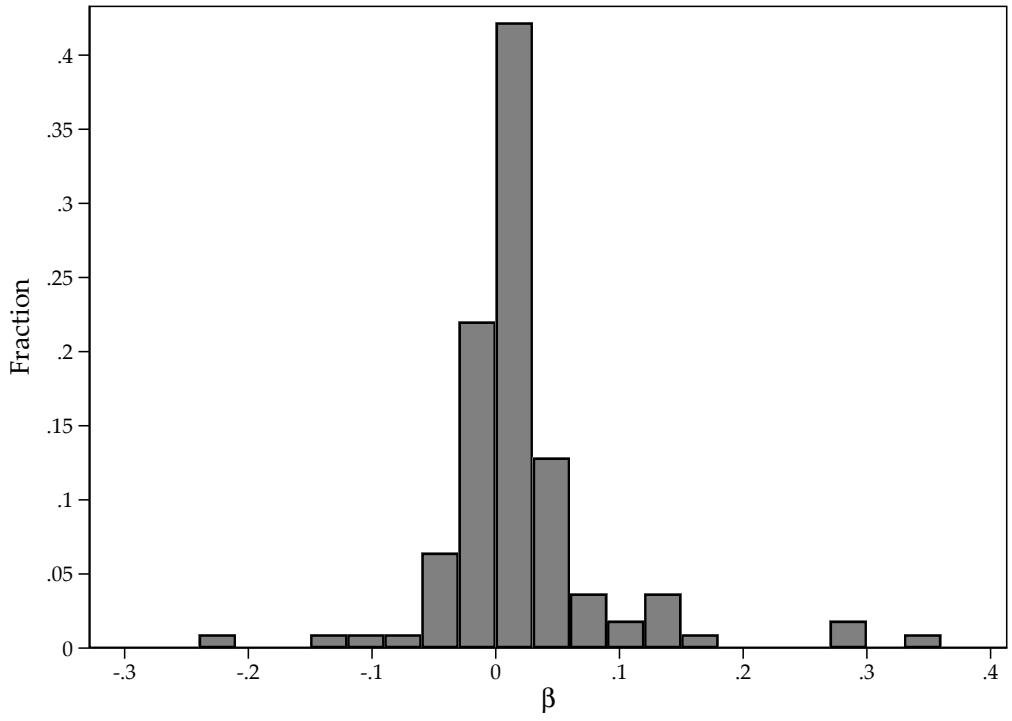



Figure 13: Distribution of the coefficient on $N_{jkt}$ from equation 4

where $N_{ijt}$ is the number of competitors in the same therapeutic class as product $j$. In the cross-section of all products in our main sample we find that $N_{jkt}$ is associate with a small but significant increase in price ($\beta = 0.0086$, $SE = 0.0008$). However, this effect may be in part due to a selection effect in the data: therapeutic classes with higher demand will have more products, and these products will also have higher prices. To partially correct for this effect, we separately run regression 4 on each therapeutic class. Figure 13 plots a histogram of the coefficients obtained on $N_{jkt}$. The distribution is centered around 0 and has a mean of approximately 0.02, which broadly confirms the results of the main regression, and suggests Bertrand is not the right pricing model for our data.

### C.7 Additional Simulation-Based Evidence of Optimality of Delays

In this section, we present additional evidence from simulation strategies. First, we examine delays in head-to-head launch strategies. In this set of simulations, we consider drugs approved between 2003 and 2008, and test strategies that limit firms to launching in two countries at most. Doing so allows us to display graphically where the impact of reference pricing is coming from. The heatmap in Figure 14 shows the fraction of drugs for which delaying in the country on the $x$-axis is optimal when the country in the $y$-axis is the only other country in the choice set. The main thing we notice is that delays are only optimal when a small number of Western European countries are in the choice set. These Western European countries reference Eastern Europe. Most

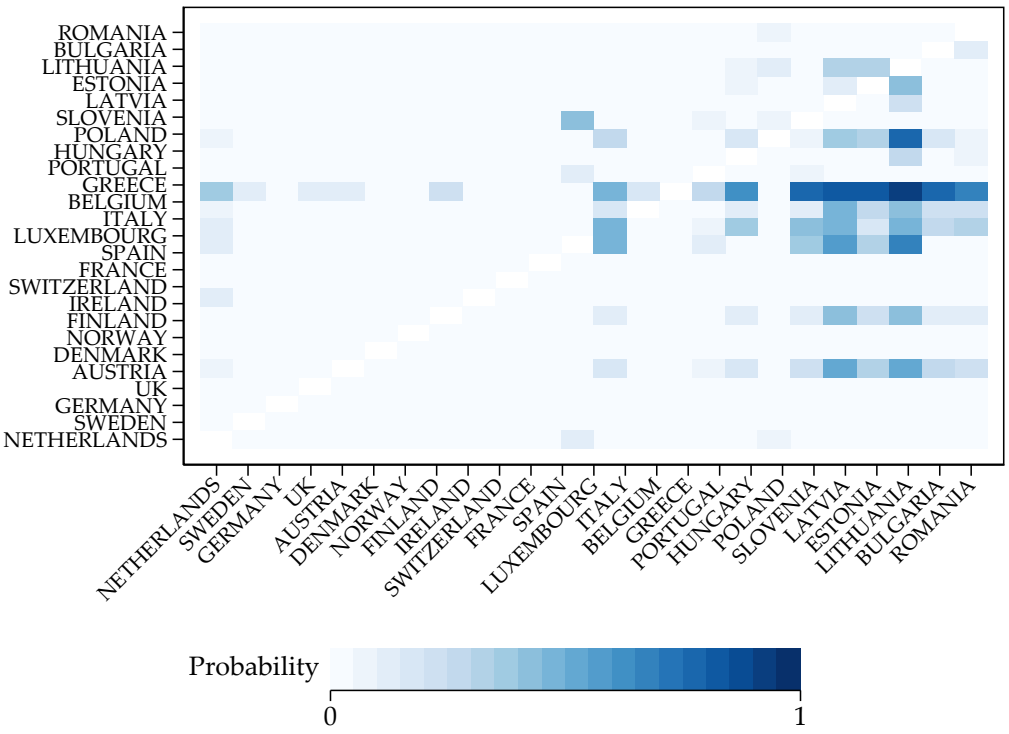



Figure 14: Likelihood of delays in head-to-head launch sequences

Western European countries either (i) do not use reference pricing, or (ii) do not reference the Eastern European country in question, meaning that they are not actually contributing to delays, at least directly.[14]

Next, we simulate delays in Eastern Europe as a block. We do this for all drugs and plot the distribution of the optimal launch date in Eastern Europe (we assume simultaneous and immediate entry in Western Europe). We find that delaying in the entire Eastern European block is optimal for almost 30 percent of drugs. The fact that this number is lower than the likelihood of delays in some smaller Eastern European countries suggest that firms can earn more by picking where to introduce the product first. This matches the launch sequences we observe in the data, where it is common for one or two Eastern European countries to be included in the first round of launches.

# D Additional theoretical derivations and extensions

In this section, we include some additional theoretical background and details for the structural estimation. We provide

1. a derivation of the nested logit demand equation

[14] They still contribute indirectly, by referencing the price of the Western European countries that directly reference Eastern Europe.

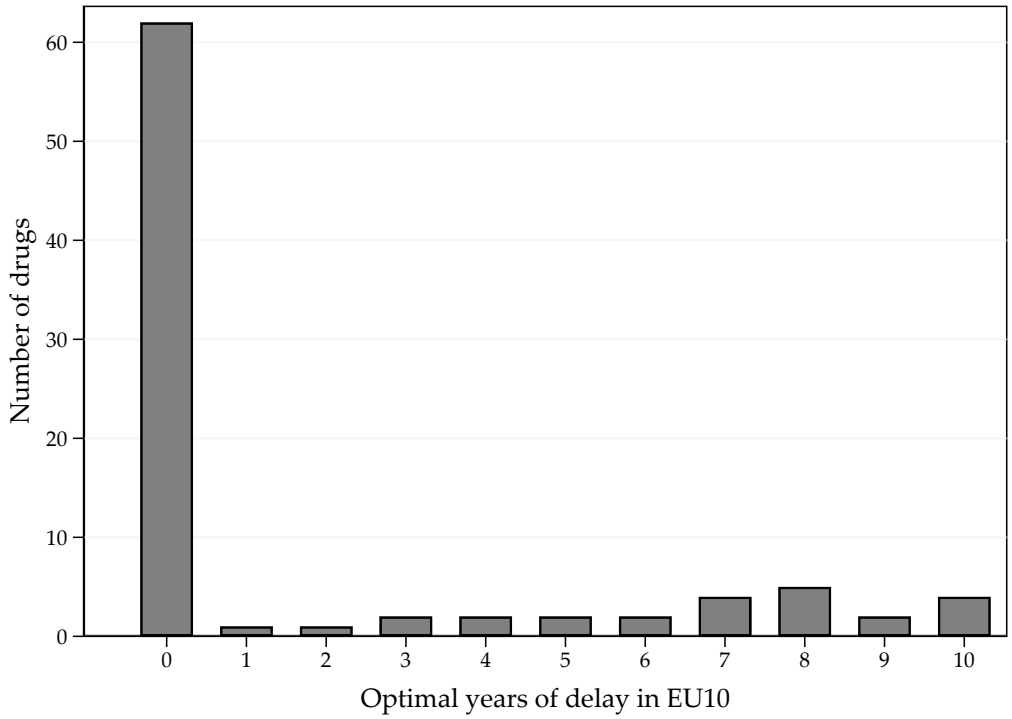



Figure 15: Distribution of optimal delay in Eastern Europe

2. a potential micro-foundation for the reduced-form pricing equation using a Nash bargaining model

3. a full description of the simulation procedure used to estimate the expected revenue of a pre-specified launch strategy for a given value of the parameter $\psi_k$ (i.e. $R_t\left(S_j, S_{-j}, X_j\right)$ from equation 12 in the main paper).

4. a derivation of an additional set of moment inequalities that can yield an upper bound on the identified set

5. a derivation of a more stringent upper bound for the entry parameter in the moment inequality

6. a full description of the procedure used to calculate the total expected revenue of a product

## D.1 Logit Model

The utility of consumer $i$, in country $k$, from consuming drug $j$ (molecule $m$) in year $t$ is given by

$$u_{ijt} = \delta_{jkt} + \zeta_{m(i)} + (1-\sigma)\,\varepsilon_{ijt} \tag{5}$$

where

$$\delta_{jkt} = \alpha_{jk} + \beta_j \text{age}_{jt} + \eta_j NF_{jkt} + \xi_{jkt}. \tag{6}$$

and $\zeta_{m(i)}$, is common for all $j \in m$, and distributed according to the unique distribution such that if $\varepsilon_{ijt}$ is an extreme value random variable, then so is $\zeta_{m(i)} + (1-\sigma)\,\varepsilon_{ijt}$ (Cardell, 1997). $\delta_{jkt}$ is parameterized as in Equation 2 in the main paper.

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

$$MS^m_{jkt} = \frac{\exp\left(\frac{\delta_{jkt}}{1-\sigma}\right)}{D_m\left(\mathbf{X}_{mt}\right)} \tag{7}$$

where

$$D_m\left(\mathbf{X}_{mt}\right) = \sum_{\ell \in m} \exp\left(\frac{\delta_{\ell jt}}{1-\sigma}\right)$$

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

$$MS_{m/kt} = \frac{D_m\left(\mathbf{X}_{mt}\right)^{1-\sigma}}{1 + \sum\limits_{h \in G} D_h\left(\mathbf{X}_{ht}\right)^{(1-\sigma)}} \tag{8}$$

where $G$ is the set of all molecules. Hence, the *overall market share of drug* $j$ is

$$MS_{jkt} = \frac{\exp\left(\frac{\delta_{jkt}}{1-\sigma}\right) D_m\left(\mathbf{X}_{mt}\right)^{-\sigma}}{1 + \sum\limits_{h \in G} D_h\left(\mathbf{X}_{ht}\right)^{(1-\sigma)}} \tag{9}$$

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

$$MS_{0kt} = \frac{1}{1 + \sum\limits_{h \in G} D_h\left(\mathbf{X}_{ht}\right)^{(1-\sigma)}} \tag{10}$$

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

$$\ln\left(MS_{jkt}\right) - \ln\left(MS_{0kt}\right) \quad = \quad \left(\frac{\delta_{jkt}}{1-\sigma}\right) - \sigma \ln\left(D_m\left(\mathbf{X}_{mt}\right)\right)$$

Combining equations 8 and 10 we also obtain

$$\ln\left(D_m\left(\mathbf{X}_{mt}\right)\right) = \frac{\ln\left(MS_{m/kt}\right) - \ln\left(MS_{0kt}\right)}{1-\sigma}$$

Hence we can write

$$\ln\left(MS_{jkt}\right) - \ln\left(MS_{0kt}\right) = \frac{\delta_{jkt}}{1-\sigma} - \frac{\sigma}{1-\sigma}\left(\ln\left(MS_{m/kt}\right) - \ln\left(MS_{0kt}\right)\right)$$

which implies

$$\begin{aligned}(1-\sigma)\left(\ln\left(MS_{jkt}\right)-\ln\left(MS_{0kt}\right)\right) &= \delta_{jkt}-\sigma\left(\ln\left(MS_{m/kt}\right)-\ln\left(MS_{0kt}\right)\right)\\ \implies \ln\left(MS_{jkt}\right)-\ln\left(MS_{0kt}\right) &= \delta_{jkt}+\sigma\ln\left(\frac{MS_{jkt}}{MS_{m/kt}}\right)\end{aligned} \tag{11}$$

## D.2 Foundations for the price-setting equation

In the main body of the paper, we do not provide a microfoundation for the price-setting equation from a utility- or revenue-based optimization model. These foundations do not fundamentally affect our results or counterfactual. For completeness however, we provide a possible set of assumptions that could justify an estimation equation identical to the one we use.

Assume that firms and governments play a Nash Bargaining game to set drug prices. The game is repeated every period, but 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 must 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 ramifications of the agreed-upon price.[15]

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

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

where $\Delta W_{jkt}$ represents the change in the welfare of the government from having drug $i$ available, $\Delta\Pi_{jkt}$ represents the incremental change in revenue, and $b_k$ is the bargaining power of country $k$. Notice that the interpretation of $\Delta W_{jkt}$ 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, $\Delta\Pi_{jkt}$ is simply the potential revenue in country $k$, 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

$$p^{\star}=\arg\max_{p}\left[\Delta W_{jkt}-p\right]^{b_k}\times\left[p\right]^{1-b_k}$$

where the interpretation of $\Delta W_{jkt}$ is the average change in the welfare function from obtaining an

[15] These assumptions are clearly unrealistic. However, 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 units operate in relative independence from one another.

additional unit of drug $i$. The standard Nash bargaining solution, can then be written as

$$p^{\star} = \Delta W_{jkt} \left(1 - b_k\right)$$

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

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 bargaining weight assigned to the government. The reference price $p_{jkt}^{\text{ref}}$ is calculated as described in the main body of the paper.[16] Given $p_{jkt}^{\text{ref}}$, we write the bargaining weight of the government as

$$B_{jk}\left(p_{jkt}^{\text{ref}}\right) = b_k + \left(1 - b_k\right)\mu_k \left(1 - \frac{p_{jkt}^{\text{ref}}}{p_{jkt}^{\text{gov}}}\right) \cdot \mathbb{I}_{\left\{p_{jkt}^{\text{ref}} < p_{jkt}^{\text{gov}}\right\}}$$

where $p_{jkt}^{\text{gov}} = \Delta W_{jkt}\left(1 - b_k\right)$ 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 $1 - B_{jk}\left(p_{jkt}^{\text{ref}}\right)$.

The function $B_{jk}\left(\cdot\right)$ has several attractive properties. First, it reduces to the base case whenever $p_{jkt}^{\text{ref}} < p_{jkt}^{\text{gov}}$. 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 $\mu_k \in (0, 1)$ the bargaining weight is also lying 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

$$[p]: \quad \left(\Delta W_{jkt} - p\right)^{-1+b+\mu_k - b_k\mu_k - \frac{\mu_k p_{jkt}^{\text{ref}}}{\Delta W_{jkt}}} \cdot \left(\left(1 - b_k\right)\left(1 - \mu_k\right)\Delta W_{jkt} + \mu_k p_{jkt}^{\text{ref}} - \Delta W_{jkt}\right) p^{-b_k(1-\mu_k)+\mu_k\left(1-\frac{p_{jkt}^{\text{ref}}}{\Delta W_{jkt}}\right)} = 0$$

and has three roots:

$$\begin{aligned} p_1^{\star} &= 0 \\ p_2^{\star} &= \Delta W_{jkt} \\ p_3^{\star} &= \left(1 - \mu_k\right)\left(1 - b_k\right)\Delta W_{jkt} + \mu_k p_{jkt}^{\text{ref}} \end{aligned}$$

[16] 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.

Notice that $p_{jkt}^{\text{ref}} \leq (1-b_k)\,\Delta W_{jkt}$ whenever the reference price binds. Hence, $p_1^\star < p_3^\star < p_2^\star$.

The second-order condition is given by

$$SOC: \quad p^{-1-b(1-\mu_k)-\mu_k\left(1-\frac{p_{jkt}^{\text{ref}}}{\Delta W_{jkt}}\right)} \cdot \left(\Delta W_{jkt}-p\right)^{-2+b_k+\mu_k-b_k\mu_k-\frac{\mu_k p_{jkt}^{\text{ref}}}{\Delta W_{jkt}}} \cdot \left(\Delta W_{jkt}\left(1-b_k\right)\left(1-\mu_k\right)+\mu_k p_{jkt}^{\text{ref}}\right)\left(-b_k\Delta W_{jkt}\left(1-\mu_k\right)+\mu_k\left(p_{jkt}^{\text{ref}}-\Delta W_{jkt}\right)\right)$$

and, for $p \in \left(0, \Delta W_{jkt}\right)$, is proportional to

$$\begin{aligned} SOC &\propto \left(\Delta W_{jkt}\left(1-b_k\right)\left(1-\mu_k\right)+\mu_k p_{jkt}^{\text{ref}}\right)\left(-b_k\Delta W_{jkt}\left(1-\mu_k\right)+\mu_k\left(p_{jkt}^{\text{ref}}-\Delta W_{jkt}\right)\right) \\ &\propto \left(-b_k\Delta W_{jkt}\left(1-\mu_k\right)+\mu_k\left(p_{jkt}^{\text{ref}}-\Delta W_{jkt}\right)\right) < 0 \end{aligned}$$

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:

$$p_{jkt} = \begin{cases} \left(1-b_k\right)\Delta W_{jkt} & \text{if } p_{jkt}^{\text{ref}} \geq \left(1-b_k\right)\Delta W_{jkt} \\ \left(1-\mu_k\right)\left(1-b_k\right)\Delta W_{jkt}+\mu_k p_{jkt}^{\text{ref}} & \text{if } p_{jkt}^{\text{ref}} < \left(1-b_k\right)\Delta W_{jkt} \end{cases}$$

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

■

### D.3 Construction of predicted reference prices in the price estimation

We briefly discuss how we construct predicted reference prices in the price estimation.

Our model states that

$$p_{jkt}\left(S_t, D_{jkt}\left(\cdot\right)\right) = \begin{cases} p_{jkt}^{\text{gov}}\left(S_t, D_{jkt}\left(\cdot\right)\right) & \text{if } p_{jkt}^{\text{ref}}\left(\cdot\right) \geq p_{jkt}^{\text{gov}}\left(\cdot\right) \\ \left(1-\mu_k\right)p_{jkt}^{\text{gov}}\left(S_t, D_{jkt}\left(\cdot\right)\right)+\mu_k p_{jkt}^{\text{ref}}\left(S_t, D_{jkt}\left(\cdot\right)\right) & \text{if } p_{jkt}^{\text{ref}}\left(\cdot\right) < p_{jkt}^{\text{gov}}\left(\cdot\right) \end{cases} . \quad (12)$$

where

$$p_{jkt}^{\text{gov}}\left(D_{jkt}\left(S_t, \xi_{kt}\right)\right) = \theta_j \cdot \gamma_k \cdot \exp\left(\beta_Z Z_{jkt} + \beta_D \ln\left(D_{jkt}\left(S_t, \xi_{kt}\right)\right)\right)$$

and $p_{jkt}^{\text{ref}}\left(S_t, D_{jkt}\left(\cdot\right)\right)$ comes from the reference pricing rules.

For drugs launched in 2002 or later, we have all the data we need to calculate predicted prices. We start from the year of launch $t_0$, and calculate $p_{jkt_0}^{\text{gov}}$ using data on data on our controls $Z_{jkt_0}$ and

$D_{jkt_0}$, and our guess for the parameters $\gamma_k$, $\mu_k$, $\beta_Z$, and $\beta_D$. We can then derive $p_{jkt_0}$ because in the first period, $p_{jkt_0} = p^{\text{gov}}_{jkt_0}$: no references are available yet. Then, we use $p_{jkt_0}$ to build the reference prices for the following period: $p^{\text{ref}}_{jkt_0+1}$. We can also build $p^{\text{gov}}_{jkt_0+1}$, again, relying only on data. With reference prices and data, we then build $p_{jkt_0+1}$, and repeat the process as long as needed.

For drugs launched before 2002, the process is only slightly different. We still observe all components of $Z_{jkt}$ for pre-2002 years in our data, i.e. home-country indicator, and number of competing molecules in class.[17] The only information we are missing is $D_{jkt}$. However, our model assumes that volume adjustments to price are excluded from the reference price.[18] This allows us to build reference prices for period $t_0 + 1$ even without an exact estimate for $p^{\text{gov}}_{jkt_0}$. Hence, for these products, our algorithm runs an initial "adjustment" loop for years 1995-2001 using only $Z_{jkt}$ and our parameter guesses.

### D.4 Simulation procedure

To simulate the function $\tilde{R}_t\left(S_j, S_{-j}, X_j\right)$ we use the following procedure:

1. Define the inputs of the functions as:

   - $\mathcal{A}'_{jt} = \{a_{k\tau}\} \rightarrow$ an array with dimension $N_C \times (T - t) + 1$ which specifies a conditional binary action for each country $k$ and period $\tau \in \{t, t+1, \ldots, T\}$: 0 means not sending an application, and 1 means sending an application (the action is conditional because it can only be taken if the application has not been sent yet). This strategy is specified by the econometrician as a guess.
   - $S_{t-1} = \left\{S_{jt-1}, S_{-jt-1}\right\} \rightarrow$ an array of dimension $N_C \times N_l$, where $N_C = 25$ is the number of EEA countries in the simulation, and $N_l$ is the number of firms in the market. This array indicates the countries where each product has launched at the beginning of the simulation, and comes from the data.
   - $\psi' = \left\{\psi'_k\right\} \rightarrow$ an $N_C \times 1$ vector whose elements all lay on the unit interval. This is the parameter vector that determines idiosyncratic delays and is specified as a guess by the econometrician.

2. Simulate an array $\{\omega_{k\tau n}\}$ of uniform random draws from the unit interval. The array has dimension $N_C \times T \times n_{sim}$ where $T = 11$ is the maximum number of years, and $n_{sim} = 500$ is the number of simulations.[19]

3. For each simulation $n$, calculate the simulated entry sequence by combining $\mathcal{A}'_{jt}$ with the array of uniform random draws: entry in country $k$ occurs in period $\tau$ if $\tau$ is the earliest

[17] We may miss molecules that exit before 2002. However, this should be a rare occurrence.

[18] As discussed in the main body of the paper, we believe this adjustment better reflect the timing of ERP application (which should occur at the beginning of the period) vs. the timing of volume adjustments (which can only occur once volume is observed, at the end of the period).

[19] In various runs we set $n_{sim}$ as high as 5,000. The associated efficiency gain was small relative to the computational gain of using only 500 simulations, so we opted to set $n_{sim} = 500$ in the main and final run.

period that satisfies $a_{kt} = 1$ and $\omega_{k\tau n} \leq \psi'_k$.

4. For each simulation $n$, calculate predicted demand and prices using the estimates of demand and price and the simulated entry sequence. In these calculations, the entry sequences of all other products are held constant.

5. Calculate total revenue by summing up revenue in each period and country where the product was available, according to the simulation. $\tilde{R}_t\left(S_j, S_{-j}, X_j\right)$ is the average of total revenue across all simulations.

### D.5 Proof that the moment inequality estimator yields a one-directional bound

**Proposition 1.** *Assume that the expected revenue conditional on playing the optimal strategy is monotonically decreasing (increasing) in $\psi$ (i.e $\frac{\partial \mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}^{\star}_{jt}\left(\mathcal{I}_{jt}\right), Y_j, \psi_0\right) \middle| \mathcal{I}_{jt}\right]}{\partial \psi} \leq (\geq)\, 0$). Then, it is impossible to find an alternative strategy $\mathcal{A}'_{jt}$ such that*

$$\mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}'_{jt}, Y_j, \psi'\right) \middle| \mathcal{I}_{jt}\right] > \mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}^{\star}_{jt}\left(\mathcal{I}_{jt}\right), Y_j, \psi_0\right) \middle| \mathcal{I}_{jt}\right]$$

*for $\psi' \geq (\leq)\, \psi^0$.*

*Proof.* The proof of this proposition is fairly straightforward. By assumption, we know that

$$\mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}^{\star}_{jt}\left(\mathcal{I}_{jt}\right), Y_j, \psi_0\right) \middle| \mathcal{I}_{jt}\right] \geq \mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}^{\star}_{jt}\left(\mathcal{I}_{jt}\right), Y_j, \psi'\right) \middle| \mathcal{I}_{jt}\right]$$

for all $\psi' \geq \psi^0$. Moreover,

$$\mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}^{\star}_{jt}\left(\mathcal{I}_{jt}\right), Y_j, \psi'\right) \middle| \mathcal{I}_{jt}\right] \geq \mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}'_{jt}, Y_j, \psi'\right) \middle| \mathcal{I}_{jt}\right]$$

for all $\psi'$, by definition. Combining the two inequalities gives the result of the proposition. □

Intuitively, the premise of the proposition holds true in our data: 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.

### D.6 Additional moment inequalities with revenue data

In the empirical implementation of the revenue-based moment inequalities we were able to generate a lower bound on the probability of an idiosyncratic delay, but not an upper bound. In this section we discuss an extension of the moment inequalities which can be helpful in generating an

upper bound in these situations.[20] The idea of this extension is to prove that for a given value of the idiosyncratic delay a the properties of the moment inequality estimator parameter no strategy will ever yield expected profits as high as what the firm obtained in the data.

Suppose that in the data we can recover a consistent estimate for $\mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}_{jt}^{\star}\left(\mathcal{I}_{jt}\right), Y_j, \psi_0\right)\middle|\mathcal{I}_{jt}\right]$ (i.e. the expected profits of the firm when playing the optimal strategy $\mathcal{A}_{jt}^{\star}\left(\mathcal{I}_{jt}\right)$ and for the true value of the parameter $\psi^0$). Recall that in the main body of the paper we defined the identified set of $\psi$ as

$$\Psi^I = \left\{\psi' : \sum_{\mathcal{A}_t'} IR\left(\mathcal{A}_{jt}, \psi'\right) = 0\right\}$$

Let $F^{ub}\left(\psi, \cdot\right)$ be a function of $\psi$ such that

$$F^{ub}\left(\psi, \cdot\right) \geq \mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}_{jt}^{\star}\left(\mathcal{I}_{jt}\right), Y_j, \psi\right)\middle|\mathcal{I}_{jt}\right]$$

for all $\psi$. Then, if $\psi'$ is such that

$$F^{ub}\left(\psi', \cdot\right) < \mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}_{jt}^{\star}\left(\mathcal{I}_{jt}\right), Y_j, \psi_0\right)\middle|\mathcal{I}_{jt}\right]$$

$\psi' \notin \Psi^I$.

To see why, suppose $\psi'$ is the true value of $\psi$, i.e. $\psi' = \psi^0$. Then

$$\mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}_{jt}^{\star}\left(\mathcal{I}_{jt}\right), Y_j, \psi_0\right)\middle|\mathcal{I}_{jt}\right] > F^{ub}\left(\psi^0, \cdot\right) \geq \mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}_{jt}^{\star}\left(\mathcal{I}_{jt}\right), Y_j, \psi_0\right)\middle|\mathcal{I}_{jt}\right]$$

Hence, $\mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}_{jt}^{\star}\left(\mathcal{I}_{jt}\right), Y_j, \psi_0\right)\middle|\mathcal{I}_{jt}\right] > \mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}_{jt}^{\star}\left(\mathcal{I}_{jt}\right), Y_j, \psi_0\right)\middle|\mathcal{I}_{jt}\right]$, which is a contradiction.

As long as $\mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}_{jt}^{\star}\left(\mathcal{I}_{jt}\right), Y_j, \psi_0\right)\middle|\mathcal{I}_{jt}\right]$ is bounded, finding candidates for $F^{ub}\left(\cdot\right)$ will not be hard. However, it may be challenging to find a function whose value is low enough to obtain meaningful inference from the data. Here we propose a possible strategy to come up with a family of such functions that could prove useful in this regard.

The basic idea is to use the expected profit of the firm when facing a simplified environment that is tweaked to be more favorable to the firm (relative to the true environment). For example, one obvious candidate for $F^{ub}\left(\cdot\right)$ in the empirical application we consider is the expected profits of the firm when ERP is eliminated. In this scenario, the firm always earns more money. Moreover, with this structure, the model has a clear, unique solution: applying for entry everywhere right away.

This function satisfies two main criteria: it is easy to compute, and it is always an upper bound for $\mathbb{E}\left[\tilde{V}_t\left(\mathcal{A}_{jt}^{\star}\left(\mathcal{I}_{jt}\right), Y_j, \psi_0\right)\middle|\mathcal{I}_{jt}\right]$. Unfortunately, this function also does not deliver a meaningful bound when tested empirically. When we tested this approach empirically we were unable to reject any value of the parameters that had not already been rejected using our original approval

[20] The same methodology can be used to obtain a lower bound in situations where the one-directional bound is the upper bound.

and entry moment inequalities. Nonetheless, this approach could provide a promising avenue for future research.

### D.7 Additional moment inequalities with entry and approval data

To derive a lower bound using entry data we calculate the overall probability of a delay across all drugs in the sample. The assumption we use in the estimation is that the probability of a delay is country-specific, but homogeneous across all drugs. Under this assumption, the overall probability of a delay in any given subsample of drugs should be higher than the probability of an idiosyncratic delay (as long as the subsample was not selected based on entry data). This suggests that we could obtain a tighter upper bound by simply looking at moment inequalities based on subsamples of the data instead of only looking at the full sample.

Formally, let $G$ be a partition of our set of drugs. For each set $g \in G$, let $\left(1-\bar{\psi}_{gk}\right)$ be the probability that product $j \in g$ will enter in country $j$ in a given year. Since by assumption $\psi_k$ is constant across drugs, it follows that $\bar{\psi}_{gk} \geq \psi_k$ for all $g$, and therefore

$$\psi_k \leq \psi_{Gj}^{\min} = \min_{g \in G} \left\{\bar{\psi}_{gk}\right\}$$

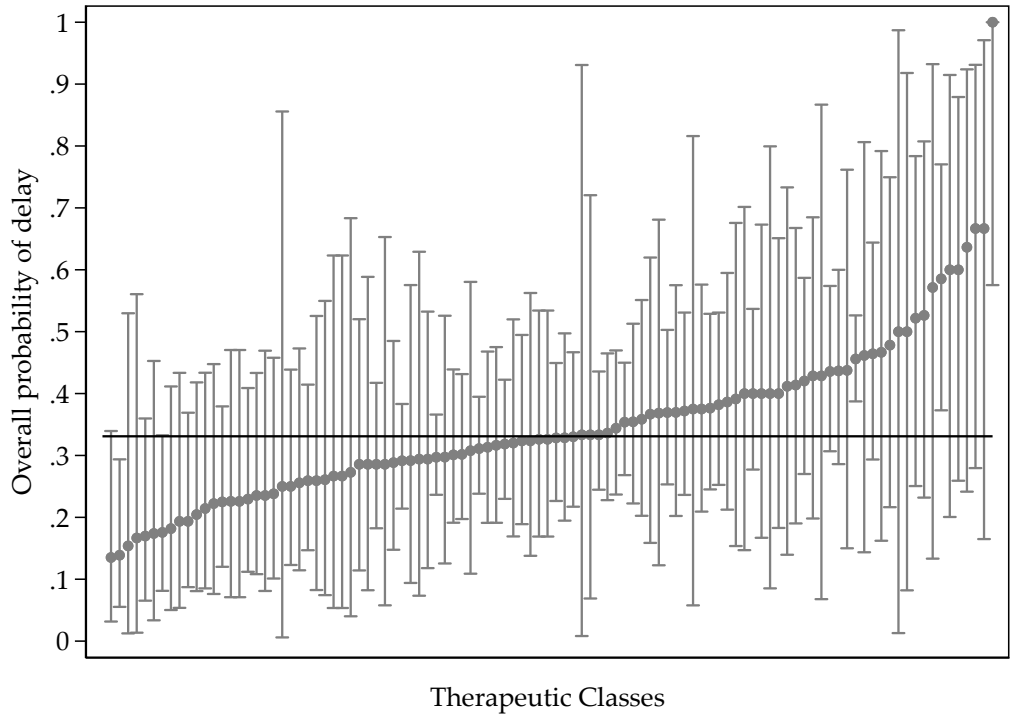
To illustrate how this method could yield a lower upper bound, we calculate the class-specific probability of delay in Eastern European countries (we aggregate across Eastern European countries to increase the sample size in each class, which would otherwise be very small in many classes). Figure 16 plots the values and confidence intervals for $\psi_{g,EE}$ for the partition of the sample generated by the therapeutic class variable. Since we may be concerned about false positives (the partition includes around 100 sets), we use 99% confidence intervals around the point estimates. We find that in three instances the parameter estimate for $\bar{\psi}_{g,EE}$ is lower and significantly different from the estimate obtained from the full sample (which is $\psi_{EE} = 0.34$). The three classes are erectile dysfunction, cytotoxics (i.e. drugs for chemotherapy), and nasal corticosteroids without anti-infectives. Using either one of these three estimates would yield a lower upper bound.

### D.8 Calculation of expected revenue loss

To calculate the expected revenue loss of a firm that chooses to disregard the impact of ERP, we simulate the expected revenue of the naive strategy (i.e. apply in all countries right after receiving marketing authorization) for each drug and compare it to the realized revenue in the data. In order to eliminate the expectational error in the realized revenue we have to aggregate across drugs, just like we did in constructing the moment inequalities.

The procedure is almost identical to the procedure we follow in computing the moment inequalities, but with one caveat, which is that in the data we do not observe the full life cycle of each drug. To address this limitation we calculate age-specific average expected revenues and then sum them up.

Figure 16: Probability of Delay in Eastern Europe by Therapeutic Class




This graph plots the overall probability of delay in Eastern European countries for drugs separated according to their therapeutic class. Vertical bars represent the 99% confidence interval, while the black horizontal line is the estimate for $\psi_{EU10}$ obtained from the overall sample.

**Calculation of expected revenue of "naive" strategy:** These are the steps we follow to calculate the expected revenue of the "naive" strategy:

1. Fix $\psi'_k = 0.69$: this is the lower bound for $\psi_{EE}$ that we estimated using the revenue inequalities.

2. For each drug $i$ in the main sample (i.e. the full sample of 481 drugs) repeat these steps:

   (a) Take the simulated array $\{\omega_{k\tau n}\}$ from the moment inequality estimation.

      i. For each simulation $n$, calculate the simulated entry sequence by assuming that firms apply right away in all countries: entry in country $k$ occurs in the earliest period $\tau$ if $\tau$ such that and $\omega_{k\tau n} \leq \psi'_k = 0.69$.

   (b) For each simulation $n$, calculate predicted demand and prices using the estimates of demand and price and the simulated entry sequence. In these calculations, the entry sequences of all other products are held constant.

   (c) Calculate age $a$ revenue $R^{\text{sim}}_{j,a}$ by summing up revenue in each country where the product was available in year $a$ after the authorization date. For years *after* a product's loss of exclusivity, we impute $R^{\text{sim}}_{j,a} = 0$.

3. We then calculate the age $a$ expected revenue of the average drug as

$$R_a^{\text{sim}} = \frac{1}{N_j} \sum_{j=1}^{N_i} R_{j,a}^{\text{sim}}$$

4. Finally, the total expected revenue of the average drug is simply $R^{\text{sim}} = \sum_a R_a^{\text{sim}}$.

**Calculation of expected revenue of optimal strategy:** These are the steps we follow to calculate the expected revenue of the optimal strategy:

1. For each drug $j$ in the main sample (i.e. the full sample of 481 drugs) calculate realized age $a$ revenue $R_{j,a}^{\text{real}}$ by summing up revenue in each country where the product was available in year $a$ after the authorization date. For years *after* a product's loss of exclusivity, we impute $R_{j,a}^{\text{sim}} = 0$.

2. We then calculate the age $a$ expected revenue of the average drug as

$$R_a^{\text{real}} = \frac{1}{N_j} \sum_{j=1}^{N_i} R_{j,a}^{\text{real}}$$

3. Finally, the total expected revenue of the average drug is simply $R^{\text{real}} = \sum_a R_a^{\text{real}}$.

We obtain the figure from the main paper (€18 million) by computing $R^{\text{real}} - R^{\text{sim}}$.

## References


**Berndt, Ernst R and Joseph P Newhouse**, "Pricing and Reimbursement in U.S. Pharmaceutical Markets," 2010.

**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.

**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," *The Quarterly Journal of Economics*, may 2019, *134* (2), 843–894.

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

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

**Feng, Josh**, "Pricing Intermediaries in Prescription Drug Markets: To Leverage or Replace?," 2020.

**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.

**Leopold, Christine, 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.

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