Demand estimation and merger simulations for drugs: logits v. aids

Farasat A.S. Bokhari
and
Franco Mariuzzo
Centre for Competition Policy
School of Economics
University of East Anglia

CCP Working Paper 16-6 v3

Abstract
We used ADHD drugs sales data from 2000-2003 and compare estimates of elasticities and merger simulations from three different demand models. Models include logit, random coefficients logit, and conditional AIDS demand model with multistage budgeting. The magnitude of cross-price elasticities is large in the third model in comparison to the first two, and some of the cross-price elasticities are estimated to be negative. Hypothetical merger simulations show large price effects for the multistage AIDS model in comparison to the discrete choice models.

Contact Details:
Farasat Bokhari f.bokhari@uea.ac.uk

The authors wish to acknowledge the support from the Centre for Competition Policy

ISSN 1745-9648

Demand estimation and merger simulations for drugs: Logits v. AIDS †

Farasat A.S. Bokhari *, Franco Mariuzzo

School of Economics, University of East Anglia, Norwich NR4 7TJ, UK

Centre for Competition Policy, University of East Anglia, Norwich NR4 7TJ, UK

Revised: January 2018

Abstract

We use ADHD drugs sales data from 2000-2003 and compare estimates of elasticities and merger simulations from three different demand models. Models include logit, random coefficients logit, and conditional AIDS demand model with multistage budgeting. The magnitude of cross-price elasticities is larger in the third model in comparison to the first two, and some of the cross-price elasticities are estimated to be negative. Hypothetical merger simulations show larger price effects for the multistage AIDS model in comparison to the discrete choice models.

Key words: Demand systems, AIDS demand, logit, random coefficients logit, discrete choice, merger simulations, psychostimulant drugs

JEL Classification: L41, K21, I11

†The authors wish to acknowledge support from Centre for Competition Policy. This draft has also benefitted greatly from comments from Gautam Gowrisankaran, Marc Ivaldi, Laura Lasio, Morten Sæthre, Otto Toivanen, Jean-Marc Zogheib, anonymous referees, as well as participants at the CCP seminar series and the annual CRESSE, BECCLE, and EARIE conferences. The usual caveats apply

*Corresponding Author: Farasat Bokhari (f.bokhari@uea.ac.uk).

1. Introduction

Competition authorities are increasingly supplementing traditional merger analysis with quantitative methods using demand models and merger simulations. In principle, the analysis is a straightforward three-step process: use pre-merger data to estimate demand parameters, recover marginal costs using equilibrium conditions, and then simulate the merger via joint profit maximization of the merging parties. At the heart of this methodology is demand estimation, which is used to obtain price elasticities. The two classes of demand models, continuous/product space or discrete choice, offer tradeoffs in terms of allowing flexibility of substitution patterns and data requirements. For instance, the continuous choice models are usually easier to estimate when there are relatively few products. On the other hand, discrete choice models can handle a large number of products but are better suited when the alternatives under consideration are substitutes. These models force substitution patterns among choices to be positive, i.e., all alternatives are gross substitutes due to the assumption that a consumer can choose only one alternative.[1] By comparison, demand models estimated in the product space do not impose this restriction, but are generally estimated with less precision if there is not enough price variation and/or if there are too many products. In turn, these differences in substitution patterns may be important for estimating marginal costs and mark-ups for merger simulations. We highlight this inherent tradeoff by estimating models from both classes using sales data from a segment of drugs where some of them can be used jointly and hence exhibit complementarities.

A typical continuous choice model is the Almost Ideal Demand System (AIDS) by Deaton and Muellbauer (1980). It allows for a flexible substitution pattern but the number of parameters to be estimated increases in the square of number of products. For this reason it is sometimes estimated within the context of multistage budgeting, which helps to reduce the parameters problem (Hausman et al., 1994). In the second class of models, logit and its extensions have been popular for demand estimation with aggregate data (Ivaldi and Verboven, 2005). Since logit imposes a substitution pattern that suffers from the Independence of Irrelevant Alternatives (IIA) problem, much attention has been given in the academic literature to allow for correlations in preferences over attributes which has lead to the more general random coefficients logit model, attributed to Berry et al. (1995) (also called the 'BLP' model). However, substitutability between products is not a condition that holds in all markets. In the case of drugs, those within the same therapeutic

[1] When there are complementarities or bundling, discrete choice models on aggregate data can still be employed by removing the constraint that a consumer must buy only one product. Nevo et al. (2005) provide an example of a random coefficients logit model where individual libraries buy all journals above a threshold until they exhaust their budget. For another example of accounting for complementarities see Gentzkow (2007). Consumer choice can be modeled as A, B or A&B and similar other combinations for multiple choices when products are bundled together (see Fox and Lazzati (2016) for identifying restrictions in such a case). Minjae et al. (2017) use discrete choice methods to handle such combinations in the context of choice of cancer drugs. To model the combined choice of A&B relative to A and B individually, one either needs data on individual choices so as to construct shares of A&B relative to A or B alone, or other assignment rules, as was the case in their work.

class, but differing by active ingredient or by formulation, may in fact be jointly consumed to form a complete therapeutic regime ("cocktail"), for instance for HIV or cancer drugs. There are many situations where complementarity is the result of bundling. Armstrong (2013) shows under which conditions this can happen, and gives examples. If this is the case, then estimating a discrete choice model, even a sophisticated one that overcomes the IIA problems, may still be committing a specification error unless complementarities are accounted for.

There are some examples where discrete choice models have been used to estimate demand for drugs, but where it is possible that some patients may be using the products in combinations. Björnerstedt and Verboven (2016) estimate a discrete choice model to evaluate post-merger prices in the Swedish market for pain killers. Patients sometimes combine or stagger 6-8 hour dosages of ibuprofen (Advil or Motrin) with acetaminophen (Tylenol) to get pain relief or to keep fever down over the course of the day. Due to combination therapy such drugs may be gross complements (see Hay et al., 2008, Mayoral et al., 2000). Similarly, Weinberg (2011) estimates logit and nested logit models to compare pre- and post- merger prices in feminine protection products in the US. The acquisition was by Proctor and Gamble, which was the leading seller of sanitary pads (Whisper and Always) of Tambrand's tampon product (Tampax). In the analysis the sanitary pads and tampons are treated as (weak) substitutes and placed in different nests. However, for some part of the population it is a recognised practice to use these products together. As a final example, consider Duso et al. (2014) who estimate a nested logit model for Type II anti-diabetes drugs using aggregate sales data in Germany. This segment consists of many solo and combination active ingredients such as metformin (a biguanide), rosiglitazone (a thiazolidinediones) or metformin and rosiglitazone joint tablets (e.g., Avandamet). Indeed this possibility is explicitly recognised in the paper and accounted for by including sales from combination tablets. However physicians also prescribe these (and other) active ingredients as separate drugs to the same patient rather than only via combination tablets, and such joint purchases would not be accounted for in aggregate sales data.[2]

We use aggregate sales data on psychostimulant drugs prescribed for treating symptoms of attention deficit hyperactivity disorder (ADHD), which differ by active ingredient (molecule) and formulation (immediate release, extended release, etc.), and where the choice set consists of both gross complements and substitutes. To see why, note that ADHD is primarily an adolescent disease where children suffering from the disorder are often prescribed psychostimulant drugs to help them cope with symptoms of ADHD during school hours, as well as provide some coverage in the late afternoon/early evening. Differences in formulation imply differences in absorption rates of the drug in the body, and hence differences in the effectiveness period of the drugs; four hours, six-eight hours, or twelve hours (henceforth just 4-hr, 8-hr or 12-hr respectively) (Conner, 2006). Thus, two drugs

[2]Several other possible combinations such as pioglitazone and glimepiride or the triple combinations such as metformin, pioglitazone and exenatide do not seem to be included in the estimation either (see Abdul-Ghani et al., 2015, for examples of triple combinations).

with the same active ingredient, produced either by the same firm or by two competitors, may at first pass appear to be substitutes, they can in fact be used in combination, where an 8-hr drug can be taken in the morning before going to school, and followed up with a 4-hr formulation after school to provide coverage throughout the working day. For some patients, the 8-hr/4-hr combination would be a complement, and may be a substitute for the more expensive once a day 12-hr drug.

In this paper we estimate logit, random coefficients logit and AIDS with multistage budgeting using standard techniques for those models and then simulate three mergers based on each of the three demand models. The AIDS model is estimated with multistage budgeting using techniques highlighted in Hausman et al. (1994). We use the methodology outlined in Nevo (2000b) to estimate the logit and the random coefficients logit models. The current data was previously used in a different context in Bokhari and Fournier (2013), and as our point of entry, we use multilevel AIDS parameters from that paper. In all models, price is treated as an endogenous variable and the same set of instruments is used to correct for the endogeneity bias, so as to put all three models on equal footing. We use the 'Hausman instruments', i.e., after controlling for market fixed effects, the price of a drug from another geographic market is used as an instrument for the price in the current market.[3]

Our merger simulation results vis-à-vis differences in logit and AIDS based simulations, are similar to those observed in monte carlo studies of these models, which generally report higher simulated prices in AIDS models. However, we add to the limited evidence on the relative performance of these models using real sales data and extend the comparisons to include drugs that may be used in combination as well as to comparisons with the random coefficients demand models (for a review see Budzinski and Ruhmer (2010)). The substitution patterns, and the corresponding merger simulations from the three models give strikingly different results. Demand estimates from the random coefficients model appear to be reasonable, especially in comparison to the logit model, as they overcome the IIA restrictions. Also, (mean) own price elasticities are greater than one in magnitude. However, these results are at odds with those from the AIDS model where several drugs appear to be gross complements, and cross-price elasticities are much larger in magnitude. This complementarity result is robust to several alternative specifications of the AIDS model and does not appear to be simply due to an aggregation bias in the data. Further, the estimated complementary makes intuitive sense as it appears among very specific drugs: those within the same molecule – but differing by formulation and coverage described above – but not across different molecules.[4]

[3]The use of such instruments has been criticised in the literature, most famously by Bresnahan (1997), and certainly the estimates given here may also be subject to potential common demand side shocks criticism, but the point here is that the same price instruments are used in all three models so as to be able to compare them.

[4]An aggregation bias could arise when different members of a household purchase alternative products over the same purchase period, for instance regular pancake syrup for some members of the household and light pancake syrup for those with diabetes. While light and regular syrups may be unrelated in demand for individuals, they would appear as complements in aggregate data. In our case of ADHD drugs, it would be if either multiple household members were diagnosed with ADHD,

The rest of the paper is organized as follows. The next section reviews the relevant literature on demand estimation methods and on merger simulations. Section three describes the product market and the data. Section four outlines the model specification in the context of our data. Section five contains three types of results: selected regression coefficients, elasticity matrices, and simulated price changes under hypothetical mergers from each of the three demand models and for three different hypothetical mergers. The last section concludes.

2. Literature Review

2.1. **Demand Models.** Earlier empirical work involving demand models used product space approach, and focused on specifying demand systems that were consistent with the utility maximization problem of a representative consumer, met the requirements for aggregating over consumers, and allowed for flexible substitution patterns between products. Typical models include the Linear Expenditure System (Stone, 1954, Barten, 1964), the Rotterdam model (Theil, 1965), the Translog model (Christensen et al., 1975), and the Almost Ideal Demand System (AIDS – Deaton and Muellbauer, 1980), each with varying degree of flexibility in the substitution patterns. However, a common difficulty in product space approach is the dimensionality problem: unless a restrictive form is used, the number of demand system parameters increases in the square of the number of products and estimation requires access to a very rich data set.

To reduce the dimensionality problem, Hausman et al. (1994) relied on weakly separable preferences of a representative consumer, along with other assumptions necessary to allow for a multistage budgeting, to specify a series of demand systems at different levels. This method, which has become popular in empirical industrial organization, involves estimating a multilevel system consisting of, for instance, top, middle, and bottom level demand systems. The top level consists of an overall demand for the segment (e.g. beer) while the middle and bottom levels specify flexible forms such as AIDS or Cobb-Douglas demand equations for subgroups of products (foreign, domestic, premium beers etc.), and then for individual brands (Miller Lite, Coors, Budweiser, etc.). Some recent applications include Ellison et al. (1997), Hausman and Leonard (2002, 2005), Chaudhuri et al. (2006) and, Bokhari and Fournier (2013).

Alternatively, the distance metric approach developed in the spatial model by Pinkse et al. (2002) offers an alternative to multistage budgeting for handling a large number of parameters by specifying the cross-price coefficients as function of distance between brands in the product space. Pinkse and

and using different drug therapies, or since our data is aggregated over a year, if the same individual was using alternative drugs during different parts of the year. While drug switches do happen soon after the initial diagnosis, so as to identify the right molecule for the person vis-a-vis the effectiveness of the main molecule and any side effects, once the correct drug is identified, it is difficult to imagine why the form would be different at different points in the year for the same individual. Nonetheless, aggregation bias cannot be completely ruled out.

Slade (2004) extend the method and apply it to the UK brewing industry, while Rojas (2008) applies this method on the AIDS demand specification of the US beer industry.

While the product space is more natural in the sense that consumers have preferences over products, the characteristics space approach due to Lancaster (1966) and McFadden (1973), and later further developed by Berry (1994) in the context of aggregate demand, views products as bundles of characteristics over which consumers have well defined preferences. Consumers (indirect) utility is defined as a function of a few relevant characteristics of the product (some of which are not observed by the econometrician), price and an idiosyncratic random error term. The characteristics approach avoids the dimensionality problem since the parameters to be estimated do not increase in the square of the number of products, or even by the number of products if the dimension of relevant characteristics does not increase with the number of brands. Further when the error term in the consumer's random utility model is specified as a Type 1 extreme value that is independently and identically distributed, it leads to the famous logit model but introduces the independence of irrelevant alternatives (IIA) problem attributed to lack of correlations in consumer preferences over products.

Allowing for correlations among groups of products, or more generally among preferences over attributes, leads to the nested logit and the random coefficients logit models (Berry et al., 1995, Cardell, 1997). Early applications, as well as important extensions with aggregate sales data, include Verboven (1996), Nevo (2000a, 2001), Sudhir (2001), Ivaldi and Verboven (2005), Petrin (2002) and Grigolon and Verboven (2014) among others. Despite the estimation difficulties associated with the BLP model, such as the role of starting values in non-linear search algorithms, choice of alternative optimization methods, accuracy of contraction mapping to invert market shares, integration to obtain market shares, role of instruments, and under-identification of variance parameters for the random components, random coefficients model has become the favorite workhorse of much of modern demand estimation (see Knittel and Metaxoglou (2014), Dubé et al. (2012), Skrainka and Judd (2011), Reynaert and Verboven (2014) and Moon et al. (2014), respectively for advancements in work related to the aforementioned issues).

2.2. **Merger Simulations.** A comparison of merger simulated predictions with actual post-merger prices, as well comparisons of different demand models in merger simulations, have both been highlighted in the literature in the context of substitute goods (Crooke et al., 1999, Peters, 2006, Huang et al., 2008, Weinberg, 2011, Weinberg and Hosken, 2013, Miller et al., 2016, 2017, Björnerstedt and Verboven, 2016). In the latter group, i.e. comparisons across demand models, Crooke et al. (1999) use monte carlo methods to generate data that equate equilibrium prices, quantities and elasticities across four different demand models (AIDS, logit, linear and log-linear) and use the specified demand parameters to predict post merger prices. They find that log-linear, followed by AIDS models, gives the largest predicted post merger prices. Huang et al. (2008) also use monte carlo methods

to generate data sets that correspond to each of the four demand models above, but where data is calibrated so as to have the same elasticity matrix at equilibrium prices. Since they include logit as one of the models, which can only have positive cross-price elasticities, their calibration and model comparisons are in the context of substitute goods. In this respect they incorporate estimating the 'wrong' model, i.e., a specification error, since data generation may be from logit demand system while the estimated model may be AIDS or vice versa. The authors find that the magnitude of the bias in the estimated elasticities varies by the type of model, but that the logit model gives the best results in terms of merger predictions, even when the true data generating process matches one of the other demand models.

Similarly, Miller et al. (2016, 2017) also use monte carlo methods and replicate earlier results of model mis-specification in merger simulations, but extend the analysis to compare the predicted price effects from 'upward pricing pressure' (UPP) methodology against merger simulations results from the same four demand models listed above. In their design, AIDS, linear, and log-linear models are all calibrated to match the initial elasticity matrix generated by the logit model. They find that the UPP methodology price predictions are more accurate when demand estimation is based on logit or linear models than the other two, where UPP under-predicts compared to the AIDS or log-linear price simulations.

Finally, and most closely related to this paper, Weinberg and Hosken (2013) use real pre- and post-merger data from two product markets (breakfast syrup and motor oil), where a handful of substitute goods were available, and a merger was observed in each case. They investigate differences in predicted post-merger prices based on linear, AIDS and two variants of logit models and also compare the predictions to the actual post-merger prices. The authors report that generally the models under-predicted the change in prices for the motor oil market, where true price changes were large, but over predicted change in prices for the breakfast syrup market, where the true price changes were negligible. Moreover, they find that the magnitude of predicted price change for the AIDS model was larger than for the logit models.

3. Product Market and Data

We use retail level sales data from the U.S. for psychostimulant drugs prescribed for the treatment of attention deficit hyperactivity disorder (ADHD). These drugs are differentiated not just by the main active ingredient/molecule, but also by the specific form of the drug. Within each molecule, drugs are available as immediate-release (IR) or extended-release (ER) forms, where the main difference is in absorbtion rates, time to peak effects, and importantly, how many times a day the drug needs to be administered as some have a 4-hour effect, and typically need to be taken upto three times a day, while others can last 8 or 12 hours and require fewer dosing regimes per day.

The data set used in this analysis consists of annual sales (quantities and revenues) between 2000 and 2003 of all ADHD drugs, and was drawn from NDCHealth's proprietary Source Territory Manager files. Quantity is provided by weight (total number of pills times the strength in milligrams of the active ingredient), and revenue is the nominal dollar value received by all pharmacies in a ZIP code for a given drug from all payers (copay plus any third party payments such as via insurance). Data was aggregated up to the county level, all dollar values were converted to constant 2000 dollars using the CPI, and price was derived as revenue divided by quantity. The analysis is restricted to counties in 778 metropolitan statistical areas as zero sales were often observed in rural counties.

Table 1 lists the names, manufacturers, and classifications (molecules and forms) of all ADHD drugs that were on the market during the study period. By 2003 there were 16 branded products available in the market along with generics for many expired patents, spanning across three main molecules: methylphenidate-HCL (MPH), mixed amphetamine salts (MAS), and dextroamphetamines (DEX). Two other molecules (pemoline and provigil) are not first line ADHD drugs while another, (atomoxetine, brand name Strattera) is an FDA approved ADHD drug but is a non-stimulant and introduced in the last year of the data series. These last three molecules are categorized as 'other' (OTH) molecules in the table. Within each molecule, drugs are sub-grouped based on their forms IR, ER and XR which roughly correspond to hours coverage by a the given drug (4, 6-8 or 12 hours). For instance, a tablet of Ritalin, listed as MPH-IR, typically provides coverage of 4 hours while Concerta, also in the same molecule but different form (MPH-XR), provides 12 hours coverage.

Note that these drugs cannot be substituted on a gram-for-gram basis, especially those is different sub-groups, and hence some dosage equivalence is needed so that one can compare price per dosage across different drugs rather than price per gram. For instance, a child who is taking 100mg of Concerta over a period of time, if switched to Ritalin (MPH-IR) would have a change in dosage to 69mg over the same period, and hence it makes sense to compare price of 100mg of Concerta to price of 69mg of Ritalin rather than to 100mg of Ritalin. The dosage conversion factors to MPH-IR are given in parenthesis in front of the name of drug.[5] Using dosage conversion factor between MPH-IR and other drugs, and World Health Organization's (WHO) definition for defined daily dosage (DDD) for MPH-IR (30mg as DDD or 0.9 grams per month), we have converted the total quantity for any given drug to defined monthly dosage (DMD). Thus if total quantity of Concerta sold in a year is X grams, it becomes (.69X/.9) DMD. Similarly, average price of Concerta, which in 2003 was \$73.94 per gram (in constant 2000 dollars) becomes $73.94 \times .9/.69 = \$96.45$ per DMD (in 2000 constant dollars). Table 1 provides the average price and shares by revenue and quantity (post DMD adjustment) for two selected years.

[5]The conversion factors were compiled by Professor Steve Hinshaw (UC Berkeley) and Dr. Peter Levine (pediatrician with Kaiser Permanente of Northern California) and are similar to the grouping given in Conner (2006). We are indebted to Professor Hinshaw and Dr. Levine for providing us with the equivalence table.

TABLE 1. Average Prices and Shares

Product	Firm	2000 Price	2000 Share (Rev)	2000 Share (Qty)	2003 Price	2003 Share (Rev)	2003 Share (Qty)
MPH-IR (∼4hrs)							
1-Ritalin (1.0)	Novartis	47.99	0.081	0.056	52.41	0.009	0.010
2-Methylin (1.0)	Mallinckrodt	37.68	0.068	0.058	31.81	0.013	0.023
3-Generic MPH-IR (1.0)	19 firms	37.40	0.185	0.163	33.06	0.026	0.046
MPH-ER (∼8hrs)							
4a-Ritalin SR (0.83)	Novartis	66.56	0.032	0.016	76.44	0.003	0.002
4b-Ritalin LA (1.25)	Novartis	.			57.25	0.024	0.025
5a-Metadate CD (1.25)	Celltech	.			56.22	0.024	0.024
5b-Metadate ER (0.83)	Celltech	65.33	0.007	0.003	66.11	0.002	0.002
6-Methylin ER (0.83)	Mallinckrodt	58.41	0.004	0.002	52.29	0.006	0.007
7-Generic MPH-ER (0.83)	15 firms	52.78	0.080	0.051	48.12	0.008	0.009
MPH-XR (∼12hrs)							
8-Concerta (0.69)	Alza/Ortho-McNeil	110.33	0.047	0.014	96.45	0.261	0.158
MAS-IR (∼4hrs)							
9-Adderall (2.86)	Shire	19.86	0.311	0.497	31.87	0.029	0.053
10-Generic MAS-IR (2.86)	3 firms	.			26.52	0.076	0.164
MAS-XR (∼12hrs)							
11-Adderall XR (2.14)	Shire	.			52.58	0.238	0.259
DEX-IR (∼4hrs)							
12-Dexedrine (1.75)	Glaxo Smith Kline	27.36	0.010	0.012	34.94	0.002	0.003
13-Dextrostat (1.75)	Shire	23.59	0.018	0.024	23.25	0.002	0.004
14-Generic DEX-IR (1.75)	4 firms	.			24.62	0.004	0.009
DEX-ER (∼8hrs)							
15-Dexedrine SR (2.14)	Glaxo Smith Kline	32.08	0.062	0.063	40.25	0.007	0.010
16-Generic DEX-ER (2.14)	2 firms	.			35.14	0.011	0.018
OTH							
17a-Cylert (0.44)	Abbott Laboratories	85.89	0.023	0.009	90.06	0.002	0.001
17b-Provigil(0.28)	Cephalon, Inc	79.13	0.059	0.025	89.85	0.094	0.061
17c-Generic Pemoline (0.44)	8 firms	64.76	0.015	0.008	60.49	0.004	0.004
17d-Strattera (0.83)	Eli Lilly	.	.		83.59	0.156	0.108

Notes Prices are per unit of defined monthly dosage (DMD). The number in parenthesis in front of the name of the drugs is the conversion factor used for converting to generic MPH-IR equivalent dosage.

Due to some data limitations, in the analysis that follows we treat certain drugs as one combined product. Within each molecule-form combination, generic drugs by different manufacturers are combined into a singe drug. For instance, in 2003 there were 19 separate manufacturers for the generic version of MPH-IR (i.e., generic version of Ritalin), and we combine the sales data from these different generic producers and treat them as one product. There are two reasons why generics within a molecule-form are treated as one product. First, individual manufactures for generics are identifiable in this data series for only the last two years, and treating these products as separate would significantly limit the number of useable observations (from four years to two years). Second, within a specific market (county-year), collectively the generics have positive sales, but individually

several products have zero sales – in 2003 the average share of the 19 MPH-IR generics collectively was 2.6% but individually in any given county many had zero sales. Since price is computed as revenue divided by quantity, price of an individual generic product cannot be ascertained for that county-year. This poses a serious estimation challenge for product space based models, such as AIDS specification, where share of a product in a market is a function of not only own price, but price of *all* other products. Just to be clear, treating generics separately implies that the share equation for Ritalin would need to be specified as a function of own price, price of 15 other branded products, and price of 49 (=15+19+3+4+8) generics. If the price of even a single generic is not available for the county-year combination (since sales for it were zero in that county-year), then no data from that county-year can be used in the analysis (and may lead to sample selection problems). Combine that with the earlier mentioned limitation of total two years of identifiable generic manufactures, and estimation is nearly impossible.

Note also that this issue of unknown price when quantity/share is zero in a county-year is equally problematic in discrete choice methods. For instance, in a logit specification, share is a function of own price and not the price of other products, but a county-year market has as many 'observations' (share-price combinations) as the number of products (think of data set in 'long-form'). This means that if the price is unknown, the share-price combination is not available for that county-year, and hence the total length of the panel is varying not due to varying choice set per market, which would be fine if the specific generic was not available in the market, but is changing because no one purchased that drug even though it is available and correct estimation would require zero share with a positive known price.

For somewhat similar reasons, we treat Novartis's Ritalin SR and Ritalin LA as one drug and Celltech's ER and CD as one drug (Novartis phased out SR tablet and introduced LA capsule and Celltech used a similar strategy). Finally drugs listed as 'OTH' are also treated as one product so that the final data set used for the alternative demand estimation analysis consists of 17 different products.

4. Model Specification

4.1. **Model 1: Standard Logit/Homegenous Tastes.** The logit assumes that there is no variation in tastes across patients and hence the indirect utility for consumer n for product j in market t (county-year combination) is given by

$$\begin{aligned} & u_{njt} = \alpha_n(y_n - p_{jt}) + \mathbf{x}_{jt}\boldsymbol{\beta}_n + \xi_{jt} + \epsilon_{njt}, \text{ where} \\ & n = 1, \ldots, N, \qquad j = 0, 1, \ldots, J, \qquad t = 1, 2, \ldots, T \text{ and where} \\ & \boldsymbol{\beta}_n = \boldsymbol{\beta}, \qquad \alpha_n = \alpha, \qquad \text{for all } N. \end{aligned} \tag{1}$$

Note that homogenous tastes are due to the assumption that $\boldsymbol{\beta}_n = \boldsymbol{\beta}$ and $\alpha_n = \alpha$ for all N. In the equation above, 0 refers to the 'outside good', chosen when the patient does not purchase any of the products. The vector $\mathbf{x}_{jt}$ (of dimension $k-1$) and random variable ξ_{jt} are the observed and unobserved (to the econometrician) product characteristics that do not vary over consumers. The former consists of molecule and form of the drug (in our case $k-1=6$), while the latter is a scalar index due to pack variety, expiration date or other unobservable characteristics related to quality.[6] The utility function in (1) can be written compactly as

$$u_{njt} = \alpha y_n + \delta_{jt} + \epsilon_{njt} \tag{2}$$

where $\delta_{jt} \equiv \alpha(-p_{jt}) + \mathbf{x}_{jt}\boldsymbol{\beta} + \xi_{jt}$ is the mean utility for product j in market t. For the logit model, we assume that ϵ_{njt} are independently and identically distributed (iid) and follow a Type-1 extreme value distribution given by $F(\epsilon) = \exp(-\exp(-\epsilon))$. In this case, the market share of product j is

$$s_{jt}(\boldsymbol{\delta}_t) = \int_{\mathbb{A}_{jt}} dF(\boldsymbol{\epsilon}) = \frac{\exp(\delta_{jt})}{\sum_{l=0}^{J}\exp(\delta_{lt})} \tag{3}$$

where $\mathbb{A}_{jt}$ is the set of characteristics of individuals that choose brand j in market t. The outside option (described below) is normalized by assuming that the price and other characteristics are zero for product 0 (i.e., $u_{n0t} = \alpha y_n + \epsilon_{n0t}$ and $\delta_{0t} = 0$) and hence the share equation can be transformed so that

$$\ln(s_{jt}) - \ln(s_{0t}) = \delta_{jt} \equiv \alpha(-p_{jt}) + \mathbf{x}_{jt}\boldsymbol{\beta} + \xi_{jt} \tag{4}$$

can be estimated using linear regression methods. Given estimates of model parameters $\boldsymbol{\theta}_1 = \begin{bmatrix}\alpha & \boldsymbol{\beta}'\end{bmatrix}'$, elasticity of product j with respect to price of product k in market t can be computed as

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}}\frac{p_{kt}}{s_{jt}} = \begin{cases} -\alpha p_{jt}(1-s_{jt}) & \text{if } j = k, \\ \alpha p_{kt} s_{kt} & \text{otherwise.} \end{cases} \tag{5}$$

For markets where the outside good has a large share (so that $1 - s_{jt}$ is small), this makes the own price elasticity to be nearly proportional to the price of the product. Also, the cross-price elasticity of product j with respect to product k depends only the market share of k, and not of all other products, i.e., it exhibits the IIA property.

Outside Good. To estimate the model in (4), we need a measure of s_0, i.e., share of the outside good. In 2003, approximately 9.5% of school aged children (age groups 4-19 years) were diagnosed with ADHD, along with a much smaller percentage of adults, and of these approximately only 60% were on ADHD medication (see Bokhari and Schneider, 2011). Thus we take total consumption of all ADHD drugs in a given market, and divide it by .1*.6 times the number of school aged children in that market to compute an approximate consumption rate. Next, to compute the *potential* size

[6] There are four molecules and three forms, each of which is specified as a dummy variable. Thus, with three dummies for molecules, two for the form, and one intercept, $k-1=6$. Also, we write the dimension of vector $\mathbf{x}_{jt}$ as $k-1$ because later we subsume the price variable in the matrix $\mathbf{X}$ which then has $k=7$ columns.

of the market, we assume that all of the diagnosed children (i.e., 100% of the 10% diagnosed), plus an additional 1% of the adult population are candidates for drug therapy, and multiply the rate above with this candidate population. In terms of quantity, this gives the (potential) market size to vary from 2.01 to 2.90 times the observed total consumption in each county, with an average value of 2.23 times the total consumption. Given the potential size of the market M_t, then based on the observed values of $q_{1t}, \ldots, q_{Jt}$, the shares of the 'inside' goods $s_{1t}, \ldots s_{Jt}$ are defined relative to this value as

$$s_{jt} = q_{jt}/M_t \qquad j = 1, \ldots, J \text{ for all } t = 1, \ldots, T \tag{6}$$

and hence the share of the outside good per market is just $s_{0t} = 1 - \sum_{j=1}^{J} s_{jt}$ for all t. This sets the average outside share to 55.08% and varies between 50.33% and 65.54% across the markets. For robustness checks, we repeated the calculation above but assumed that the true prevalence of ADHD among children is 15%, i.e., an additional 5% are undiagnosed and that all of them could potentially be prescribed ADHD medication plus the 1% of adult population. This increased the average share of the outside good to 67.33% and that of the average market size to 3.06 times the total consumption.[7]

Price Instruments. The error term in (4) is likely to be correlated with the prices so that $\text{cov}(p_{jt}, \xi_{jt}) \neq 0$ and hence as a first step towards consistent estimation, we estimate a fixed effects model with dummies for individual products and markets so that the equation becomes $\ln(s_{jt}) - \ln(s_{0t}) = \delta_{jt} = \alpha(-p_{jt}) + \mathbf{x}_{jt}\boldsymbol{\beta} + \xi_j + \xi_t + \Delta\xi_{jt}$, where ξ_j is the brand fixed effect and ξ_t is the market fixed effect. Since brand characteristics (molecule and form) do not change by individual markets, those variables are dropped in favor of the brand dummies. Market fixed effects are proxied with linear and quadratic time variables and dummies for state variables.

Once brand specific dummy variables are included in the regression, the error term is just the market specific deviation from the mean of the unobserved characteristics. For instance local level

[7] We also computed the potential size and share of outside good in a very different manner but it led to such extreme values of potential market and share of outside good that we did not use this alternative definition in our final analysis but report it here briefly. We used a 12-hr (day-long) coverage of a standard dose of ADHD drug multiplied by candidate population in each area as the total potential size of the market. According to WHO, a standard dose is a 30mg pill of MPH-IR (i.e., immediate release tablet of methylphenidate such as Ritalin), and if taken three times a day it would provide a 12-hr coverage. Thus a standard dose is 90mg of MPH-IR per day per person or 10.8 grams per year. We multiplied this dose with 15% of all school aged children and 1% of adult population to compute the market size. If we reduce the population percentages, it produces potential market size to be smaller than observed quantity in some markets (there is large cross-sectional variation in actual consumption rates and indeed even in diagnosis). However keeping it at this value of population produces a share of outside good that varies from 18.55% to 99.76% with an average value of 90%, and of market size that varies between 1.23 to 417.38 with an average value of 18.59 times total consumption. Results from such an extreme potential size market definition are available upon request, but briefly, while they don't change the regression coefficients or estimates of own-price elasticities by much, they make the cross-price elasticity estimates smaller by an order of magnitude (see formula for logit price elasticities to see why). Following on from that, merge simulations also show considerably smaller impact on price changes for competitors' products.

variation in pack variety or local promotions, and may still require the use of instruments for price if $\mathrm{E}(\Delta\xi_{jt}p_{jt}|\mathbf{x}_{jt},\xi_j,\xi_t)=0$ does not hold. Accordingly, we use instruments originally due to Hausman et al. (1994). They use the panel nature of data and the assumption that prices in different areas are correlated via common cost shocks, to use a price vector from other city as instrument for the price vector in a given city. The identifying assumption is that after controlling for brand specific intercepts and demographics, the city specific valuations of a product are independent across cities but may be correlated within a city over time. Given this assumption, the price of a brand in another city is a valid instrument, so that price of brand j in two cities will be correlated due to the common marginal cost, but due to the independence assumption will be uncorrelated with the market specific valuation of the product. Since our data is at county level, we use the average price of a given drug from 20 randomly selected counties that are not from the same census region as the original county (so as to minimize the possibility of common demand side shocks).[8]

4.2. **Model 2: Random Coefficients Logit/Heterogenous Tastes.** For the heterogenous consumer case, the general set up of utility for consumer n for product j in market t is as in (1), but without imposing the restriction that taste parameters $\{\alpha,\boldsymbol{\beta}\}$ – the marginal utilities of product characteristics – are the same for all consumers. Instead each consumer is assumed to have a different set of coefficients $\{\alpha_n,\boldsymbol{\beta_n}\}$ which are modeled as a function of underlying common parameters $\{\mathbf{\Pi},\,\mathbf{\Sigma}\}$ that are multiplied to the person specific characteristics $(\boldsymbol{d}_n,\boldsymbol{\nu}_n)$. These include observable demographic information as well as other unobservable characteristics such as physiology, type of ADHD, and school circumstances (e.g. presence of a school nurse that can administer medication during school hours if a child is taking multiple doses of 4-hr drugs within a day) that may effect the choice of drug. Thus $(\boldsymbol{d}_n,\boldsymbol{\nu}_n)$ are random draws from a mean zero population with distribution functions $F_{\boldsymbol{d}}(\boldsymbol{d})$ and $F_{\boldsymbol{\nu}}(\boldsymbol{\nu})$, and

$$\begin{bmatrix}\alpha_n\\ \boldsymbol{\beta}_n\end{bmatrix}=\underbrace{\begin{bmatrix}\alpha\\ \boldsymbol{\beta}\end{bmatrix}}_{\boldsymbol{\theta}_1}+\underbrace{\mathbf{\Pi}\boldsymbol{d}_n+\mathbf{\Sigma}\boldsymbol{\nu}_n}_{\boldsymbol{\theta}_2=\{\mathbf{\Pi},\mathbf{\Sigma}\}}. \tag{7}$$

The person specific coefficients are equal to the mean value of the parameters $\boldsymbol{\theta}_1=\begin{bmatrix}\alpha & \boldsymbol{\beta}'\end{bmatrix}'$, plus deviation from the mean due to a second set of parameters $\boldsymbol{\theta}_2=\{\mathbf{\Pi},\mathbf{\Sigma}\}$ and given by $\mathbf{\Pi}\boldsymbol{d}_n+\mathbf{\Sigma}\boldsymbol{\nu}_n$. Note that $\mathbf{\Pi}$ is $k\times D$ matrix of parameters and $\mathbf{\Sigma}$ is $k\times k$ matrix of parameters.

[8] A commonly used alternative is to employ the 'BLP instruments', i.e., sums of characteristics of other products produced by the same firm or by other competing firms. Berry et al. (1995) made the argument that changes in exogenous characteristics of competing products should help identify substitution patterns. However, as explained in Berry and Haile (2014), while these instruments are useful, they are not sufficient to alone identify demand and require additional instruments, such as cost shifters or cost proxies (i.e. prices from other markets as in the case of Hausman instruments). What prevents us from using these instruments is that, as mentioned above, we have a short panel and the fact that drug characteristics are the same in all markets, and hence there is no variation in observed product characteristics across markets that can be exploited.

If we insert (7) back into (1) and simplify, then the utility function can be decomposed into three parts (or four, if we count $\alpha_n y_n$ term but that drops out later on) and can be written as

$$\begin{aligned} u_{njt} &= \alpha_n y_n + \delta_{jt} + \mu_{njt} + \epsilon_{njt} \qquad \text{where,} \\ \delta_{jt} &= \delta(\mathbf{x}_{jt}, p_{jt}, \xi_{jt}; \boldsymbol{\theta}_1) = \alpha(-p_{jt}) + \mathbf{x}_{jt}\boldsymbol{\beta} + \xi_{jt} \\ \mu_{njt} &= \mu(\mathbf{x}_{jt}, p_{jt}, \boldsymbol{d}_n, \boldsymbol{\nu}_n; \boldsymbol{\theta}_2) = (-p_{jt}, \mathbf{x}_{jt})(\boldsymbol{\Pi}\boldsymbol{d}_n + \boldsymbol{\Sigma}\boldsymbol{\nu}_n). \end{aligned} \tag{8}$$

As before, δ_{jt} is the mean utility of drug j and is common to all patients, and $\mu_{njt} + \epsilon_{njt}$ is the mean-zero heteroscedastic error term that captures the deviation from the mean utility. The market share of product j is the integral of the joint distribution of $(\boldsymbol{d}, \boldsymbol{\nu}, \boldsymbol{\epsilon})$ over the mass of individuals in region $\mathbb{A}_{jt}$,

$$s_{jt} = \int_{\mathbb{A}_{jt}} dF(\boldsymbol{d}, \boldsymbol{\nu}, \boldsymbol{\epsilon}) = \int_{\mathbb{A}_{jt}} dF_{\boldsymbol{d}}(\boldsymbol{d}) dF_{\boldsymbol{\nu}}(\boldsymbol{\nu}) dF_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}) \tag{9}$$

where the second part follows only if we assume that the three random variables for a given consumer are independently distributed. If we continue to assume that ϵ_{njt} is iid and drawn from extreme value distribution, then the probability that a given individual n – with endowed values of $\boldsymbol{d}_n$ and $\boldsymbol{\nu}_n$, or equivalently with a given value of μ_{njt} – chooses product j has the usual logit form given by $s_{njt} = \exp(\delta_{jt} + \mu_{njt}) / \sum_{j=0}^{J} \exp(\delta_{jt} + \mu_{njt})$. Integrating this individual probability over a distribution of $\boldsymbol{d}_n$ and $\boldsymbol{\nu}_n$ recovers the market share of product j.

To obtain the model predicted market shares, we use census data and randomly select 40 adults from each MSA and year and record their vector of demographics $\boldsymbol{d}_n$ (income, health insurance, family size and race (white or not)) to construct a non-parametric distribution for $F_{\boldsymbol{d}}(\boldsymbol{d})$. Next, for $F_{\boldsymbol{\nu}}(\boldsymbol{\nu})$, we use a multivariate normal distribution and numerically compute the market share via the smooth simulator. Specifically, given $N_s = 40$ random draws for $(\boldsymbol{\nu}_n, \boldsymbol{d}_n)$ for each market t, and an initial guess of mean utility and parameters $\boldsymbol{\theta}_2 = \{\boldsymbol{\Pi}, \boldsymbol{\Sigma}\}$, model predicted market share $\tilde{s}_{jt}$ is obtained as

$$\tilde{s}_{jt} = \int_{\mathbb{A}_{jt}} s_{njt} dF_{\boldsymbol{d}}(\boldsymbol{d}) dF_{\boldsymbol{\nu}}(\boldsymbol{\nu}) = \frac{1}{N_s} \sum_{n}^{N_s} s_{njt} = \frac{1}{N_s} \sum_{n}^{N_s} \Big\{ \frac{\exp(\delta_{jt} + \mu_{njt})}{\sum_{l=0}^{J} \exp(\delta_{lt} + \mu_{nlt})} \Big\}. \tag{10}$$

The model parameters $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ are estimated using Berry's two step GMM method. Briefly, given the random draw and starting values of $\boldsymbol{\theta}_2$, first we obtain an estimate of the mean utility δ_{jt} via the contraction mapping

$$\boldsymbol{\delta}_t^{h+1} = \boldsymbol{\delta}_t^{h} + [\ln(\mathbf{s}_t) - \ln(\tilde{\mathbf{s}}_t)] \tag{11}$$

that matches model predicted market shares $\tilde{s}_{jt}$ with the observed shares s_{jt} for each market. Next, we define the error term as $\boldsymbol{\xi}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \widehat{\boldsymbol{\delta}}(\mathbf{s}, \boldsymbol{\theta}_2) - \mathbf{X}\boldsymbol{\theta}_1$, and given a suitable matrix of instrumental variables $\mathbf{Z}$ (described below), search for values of $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ that minimize the GMM objective function

$$\boldsymbol{\xi}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)' \mathbf{Z} \boldsymbol{\Phi} \mathbf{Z}' \boldsymbol{\xi}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \tag{12}$$

where $\mathbf{\Phi}$ is the GMM weighting matrix. The process is then repeated by computing new values of individual and model predicted market shares. We implement the algorithm following Nevo (2000b) and the computer code given therein.[9]

The price elasticities of market shares are also numerically estimated as

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} -\frac{p_{jt}}{s_{jt}} \int_{\mathbb{A}_{jt}} \alpha_n s_{njt}(1 - s_{njt}) dF_{\boldsymbol{d}}(\boldsymbol{d}) dF_{\boldsymbol{\nu}}(\boldsymbol{\nu}) & \text{if } j = k, \\ \frac{p_{kt}}{s_{jt}} \int_{\mathbb{A}_{jt}} \alpha_n s_{njt} s_{nkt} dF_{\boldsymbol{d}}(\boldsymbol{d}) dF_{\boldsymbol{\nu}}(\boldsymbol{\nu}) & \text{otherwise.} \end{cases} \tag{13}$$

Instruments Matrix. The instrument for price, as before, is the Hausman instrument and this variable and the brand characteristics (or brand dummies if instead they are used) form the main elements of the instruments matrix $\mathbf{Z}$ as the set of exogenous variables. Note however that the exogenous brand characteristics (or brand dummies if they were part of $\mathbf{X}$) plus the one additional (Hausman) instrument for price will give exactly as many moment conditions as the number of components of the parameter vector $\boldsymbol{\theta}_1$. These would be enough in the linear logit case. However, in the random coefficients case, we have to estimate additional $k \times D + k \times k$ parameters of $\boldsymbol{\theta}_2 = \{\mathbf{\Pi}, \mathbf{\Sigma}\}$ where, in our case, $D = 4$ is the number of person specific observed characteristics, and $k = 7$ is the length of person specific shock for each observed product characteristic (i.e. the product characteristics with random coefficients).[10] Estimating these additional parameters is not possible unless we have additional $k \times D + k \times k$ moment conditions (see Moon et al., 2014). Since these additional parameters are due to the $\boldsymbol{\mu}_n(\cdot) = \mathbf{X}(\mathbf{\Pi}\boldsymbol{d}_n + \mathbf{\Sigma}\boldsymbol{\nu}_n)$ term in the utility function, and involve interactions of the product characteristics with individual specific terms $\boldsymbol{d}_n, \boldsymbol{\nu}_n$, we use the average over N_s individuals of these interaction terms as additional variables in the instruments $\mathbf{Z}$ matrix.[11]

4.3. **Model 3: Multistage Budgeting and AIDS.** We can model a multistage budgeting system for a representative consumer to estimate demand for the drugs as follows.[12] At the top level (level 4), a representative patient decides how many units of ADHD drugs to consume as a function of a price index for these drugs, her income and other demographics. At the next stage down (level 3), budget is allocated to choice of the four molecules. The next level further allocates the budget

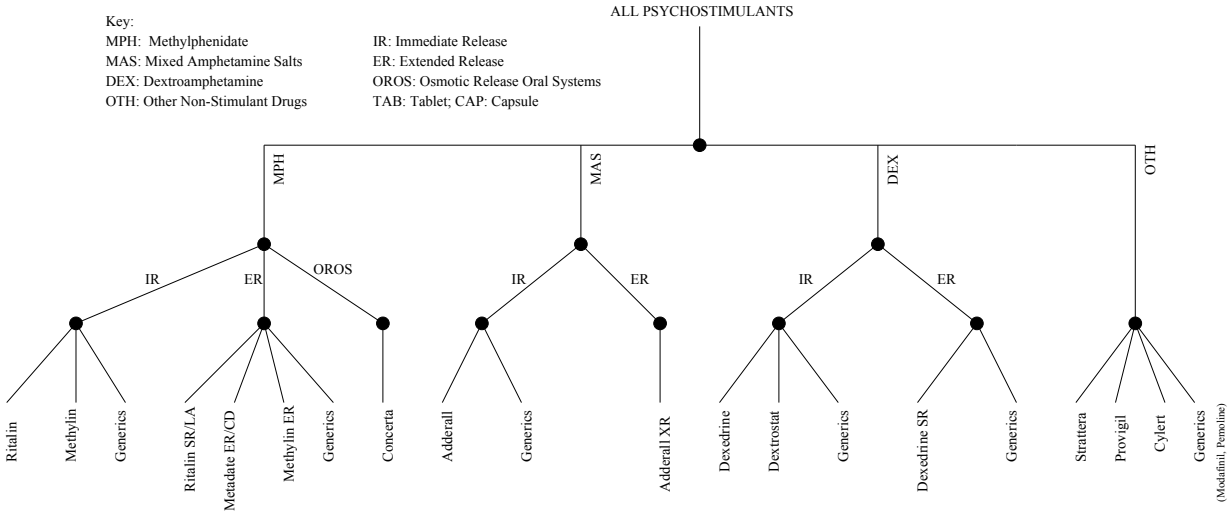
[9] As pointed out in Nevo (2000b), the matrix $\mathbf{X}$ consisting of price and product characteristics enters the model twice, once linearly via the term $\delta()$, which is common to all individuals, and a second time non-linearly via $\mu()$ that allows for random coefficients for product characteristics. Thus, $\mathbf{X}$ in each part can be written as $\mathbf{X}_1$ and $\mathbf{X}_2$, and these need not be the same set of variables. We set $\mathbf{X}_1$ as price and brand dummies, and $\mathbf{X}_2$ as price and product characteristics.

[10] In general if we set some of the terms of the $\mathbf{\Pi}$ matrix to be zero, and also set the parameter matrix $\mathbf{\Sigma}$ to be diagonal, it reduces the need for additional moment conditions from $kD + k^2$ to $g + k$ where g is the number of non-zero terms in $\mathbf{\Pi}$.

[11] An alternative would be to obtain additional moment conditions via the use of the 'optimal' instruments, i.e., expected value of derivatives of the error term with respect to the parameters vector (see equation 12 above), which in the context of the BLP instruments have been shown to improve the estimation of random coefficients models (see Reynaert and Verboven, 2014).

[12] The multilevel AIDS model with ADHD data was estimated in Bokhari and Fournier (2013), and we use the parameter estimates provided therein. Here we keep estimation details to a minimum, and refer an interested reader to the given citation.

to the choice of the forms, and at level 1 the consumer chooses among specific brands or generics within each molecule and form, conditional on the budget for the segment (see Figure 1).



Note: Generics refer to several manufacturers for each molecule and form given in the column. There are no generic versions of Concerta and Adderall XR during the study period.

FIGURE 1. Multilevel Budgeting

Specifically, if there are M molecules indexed by m, then within a molecule there are F_m forms indexed by f_m, and for a given molecule and form there are J_{fm} drugs indexed by j_{fm} such that $\sum_m^M \sum_f^{F_m} \sum_j^{J_{f_m}} j_{f_m} = J$ is the total number of drugs (since we are describing a different demand model here, some of the notation will be recycled, and should not be confused with its use in the discrete choice models discussed earlier). Then the system of demand equations for each level are

$$
\begin{aligned}
&\text{Level 4:} \quad \ln Q_t = A + B\ln(I_t) + G\ln P_t + \mathbf{x}_t\boldsymbol{\lambda} + \zeta_t \\
&\text{Level 3:} \quad \ln(Q_{tm}) = A_m + B_m \ln(y_t) + \sum_{m'}^{M} \Gamma_{mm'} \ln P_{tm'} + \mathbf{x}_{tm}\boldsymbol{\lambda}_m + \xi_{tm} \\
&\text{Level 2:} \quad u_{tf_m} = a_{f_m} + b_{f_m} \ln(\frac{y_{t_m}}{P_{t_m}}) + \sum_{f'}^{F_m} g_{ff'_m} \ln P_{tf'_m} + \mathbf{x}_{tf_m}\boldsymbol{\lambda}_{f_m} + \mu_{tf_m} \\
&\text{Level 1:} \quad s_{tj_{fm}} = \alpha_{j_{fm}} + \beta_{j_{fm}} \ln(\frac{y_{t_{fm}}}{P_{t_{fm}}}) + \sum_{j'}^{J_{fm}} \gamma_{jj'_{fm}} \ln P_{tj'_{fm}} + \mathbf{x}_{tj_{fm}}\boldsymbol{\lambda}_{j_{fm}} + \varphi_{tj_{fm}}.
\end{aligned}
\tag{14}
$$

Level 1 equations are in AIDS specification where $s_{tj_{fm}}$ is the (revenue) share of drug j in market t in segment fm, which is a function of $\ln P_{tj'_{fm}}$, the log price of all j' drugs in segment fm, as well as a function of log ratio of $y_{t_{fm}}$ and $P_{t_{fm}}$, the total expenditure and a price index for the fm segment respectively. Similarly, $\mathbf{x}_{tj_{fm}}$ and $\varphi_{tj_{fm}}$ represent other exogenous variables and the error term that effect the shares of the drug. The next level up is also an AIDS specification where u_{tf_m}

is the share of the form f in molecule m in market t. Other variables have similar interpretations where, note that the 'price' of form within a molecule (the term $\ln P_{tf'_m}$ in level two equation) is the same as the price index for that form in level 1 equation, and that the right hand side now includes a price index for the molecule (i.e., the $\ln P_{t_m}$ term). Level 3 is specified as Cobb-Douglas demand equations where 'price' of the molecule is the price index used at the previous level. Finally at the top level, log quantity of all ADHD drugs is a function of the price index for these drugs, total income and other demographic variables.

At each level the price index is constructed as the Stone price index, i.e., a share weighted average of log price rather than Deaton and Muellbauer's exact price.[13] Thus at the bottom level, the index for the price of molecule-form fm is computed as $\ln P_{t_{fm}} = \sum_j^{J_{fm}} s_{tj_{fm}} \ln P_{tj_{fm}}$.

The set of equations above also include exogenous variables that may effect shares or total quantity. These are state fixed effects, time trends (up to cubic terms), (log of) number of children, number of physicians, state level medicaid enrollees, and state medicaid expenditures on drugs. Each segment is a system of equations and is estimated separately. At the lowest two levels, homogeneity, symmetry and adding-up restrictions are imposed. In total there are five bottom level segments with two or more drugs in them. At the next level up there are three separate segments with multiple equations. For instance within the MPH segment, three equations are for IR, ER and XR while within MAS there are just two equations. At the next level there is just one segment with four equations (one for each molecule) and finally at the top level there is a single equation. The equations were estimated using OLS and 3SLS where in the later case, prices are treated as endogenous. The instrument used for prices (price from other markets) is the same as that used in the earlier specified discrete choice models.

Based on estimated parameters, cross-price elasticity between two drugs i and k in molecule-form segments f_m and $f'_{m'}$ respectively are given by

$$\begin{aligned}\frac{\partial lnQ_{i_{f_m}}}{\partial lnP_{k_{f'_{m'}}}} &= \left(1+\frac{\beta_{i_{f_m}}}{s_{i_{f_m}}}\right)\bar{s}_{k_{f'_{m'}}}\left[\frac{g_{ff'_{m'}}}{u_{fm}}+\bar{u}_{f'_{m'}}\right]\cdot\delta_m^{m'} \\ &+\left(1+\frac{\beta_{i_{f_m}}}{s_{i_{f_m}}}\right)\bar{s}_{k_{f'_{m'}}}\left[\frac{b_{fm}\bar{u}_{f'_{m'}}}{u_{fm}}+\bar{u}_{f'_{m'}}\right]\Gamma_{mm'}+\frac{1}{s_{i_{f_m}}}\left\{\gamma_{ik_{f'_{m'}}}-\beta_{i_{f_m}}\bar{s}_{k_{f'_{m'}}}\right\}.\delta_{fm}^{f'm'}-\delta_{ifm}^{kf'm'}\end{aligned} \tag{15}$$

where δ_a^b is the Kronecker delta function equal to 1 if $a = b$ and 0 otherwise.

[13] Note however that by using the Stone price index, even if the original prices were not endogenous, we are introducing an artificial endogeneity, as equation (14) now involves shares on both the left and right hand side of the equation. This is easily overcome by using period specific average value of shares from each county (average is over counties) to construct the price index for market t. To be clear, right hand side in the Stone price index uses $\bar{s}_{j_{fm}}$, i.e., period specific average share so that each time period has a different value, but is the same for all counties. Higher level price indexes are constructed the same way.

5. Results

5.1. **Regression Coefficients.** Table 2 shows selected coefficients from the discrete choice models. The first three columns show coefficients from simple logit. All three specifications include time, time square, state fixed effects and log of number of children and log of number of physicians in the area (county) as well as state medicaid expenditures on medication and total number of enrollees in medicaid (both in logs). Specifications (1) and (2) are OLS results but in (1) we use drug characteristics (molecule and form of the drug) while in (2) we instead use brand dummies. [14] Use

Table 2. Coefficients from Discrete Choice Models

	Logit			Random Coefficients Logit (4)					
	(1)	(2)	(3)	$\boldsymbol{\theta}_1$	$\boldsymbol{\theta}_2$ (Interactions with Demographic Variables, π_{kd})				
		α, β		α, β	σ_{kk}	Income	Insurance	Size	Race
Price	-2.074	-4.259	-8.755	-9.548	0.009	1.770	-7.023	0.765	-0.786
	(0.067)	(0.080)	(0.155)	(0.301)	(0.070)	(0.363)	(1.703)	(0.302)	(0.664)
MPH	-0.320	-0.418	-0.267	-0.246	-0.131	0.081	0.633	-0.186	-0.769
	(0.031)	(0.851)	(0.880)	(0.055)	(0.064)	(0.124)	(0.320)	(0.098)	(0.245)
MAS	0.838	0.740	0.936	1.616	-0.253	0.161	-2.158	-0.074	-0.754
	(0.034)	(0.852)	(0.881)	(0.064)	(0.221)	(0.143)	(0.600)	(0.107)	(0.270)
DEX	-1.542	-1.640	-1.455	1.161	-0.194	0.509	0.018	-0.179	0.452
	(0.034)	(0.852)	(0.880)	(0.058)	(0.125)	(0.144)	(0.618)	(0.101)	(0.304)
ER	-0.299	-0.299	-0.322	-1.477	-0.084	-0.420	-0.748	-0.200	0.644
	(0.016)	(0.013)	(0.014)	(0.055)	(0.094)	(0.074)	(0.390)	(0.058)	(0.146)
XR	1.351	1.351	1.265	0.125	-0.071	-0.931	0.814	-0.563	0.609
	(0.036)	(0.029)	(0.034)	(0.022)	(0.188)	(0.179)	(0.586)	(0.176)	(0.348)
Constant	-1.253	-1.253	-1.327	-3.819	-0.018	-0.478	4.103	-0.010	0.167
	(1.053)	(0.850)	(0.878)	(0.055)	(0.033)	(0.188)	(0.986)	(0.131)	(0.339)

Note: Table shows coefficients associated with price from simple logit (1,2,3) and random coefficients logit model (4). (1) is OLS without brand dummies, (2) is OLS with brand dummies, (3) is instrumental variables and brand dummies. Logits (1,2,3) also include other market specific variables and state dummies (not shown). Standard errors are in parenthesis below the coefficients. Price coefficients in this table are when price is in $100 per DMD, i.e., divided by 100, as some scaling was necessary to overcome convergence issues in estimating the random coefficients logit model.

of brand dummies nearly doubles the value of the price coefficient. Specification (3) uses brand

[14] In columns 2,3, and 4, coefficients and standard errors for brand characteristics are retrieved via GLS regression of brand dummies on brand characteristics.

dummies and instruments for price and once again the coefficient nearly doubles from its previous value. This general increase in the magnitude of the price coefficient is consistent with several previous studies that show a similar effect of including brand dummies and use of instruments (see for instance Nevo (2001), Petrin (2002)).

Specification (4) with random coefficients uses the Hausman instrument for prices and includes both the brand dummies and drug characteristics. This model shows a further small increase in the mean value of the price coefficient (-9.548). However, the coefficient on the standard deviation for price is not significant but those on the interaction term between price and income, insurance and family size are all statistically significant (though the interaction with race is not), indicating that most of the heterogeneity in price sensitivity is explained via demographics. The large dispersion in price sensitivity across individuals is shown in Figure 2. Similarly, the estimates for the standard deviations of taste parameters are not significant for any of the brand characteristics except for the dummy variable for the molecule MPH (and marginally for DEX at the 10 percent significance level). However, several interactions between characteristics and demographics are significant. Thus, relative to the other molecule, marginal utility of the MPH molecule increases with insurance but decreases with family size, with similar interpretations for other coefficients.

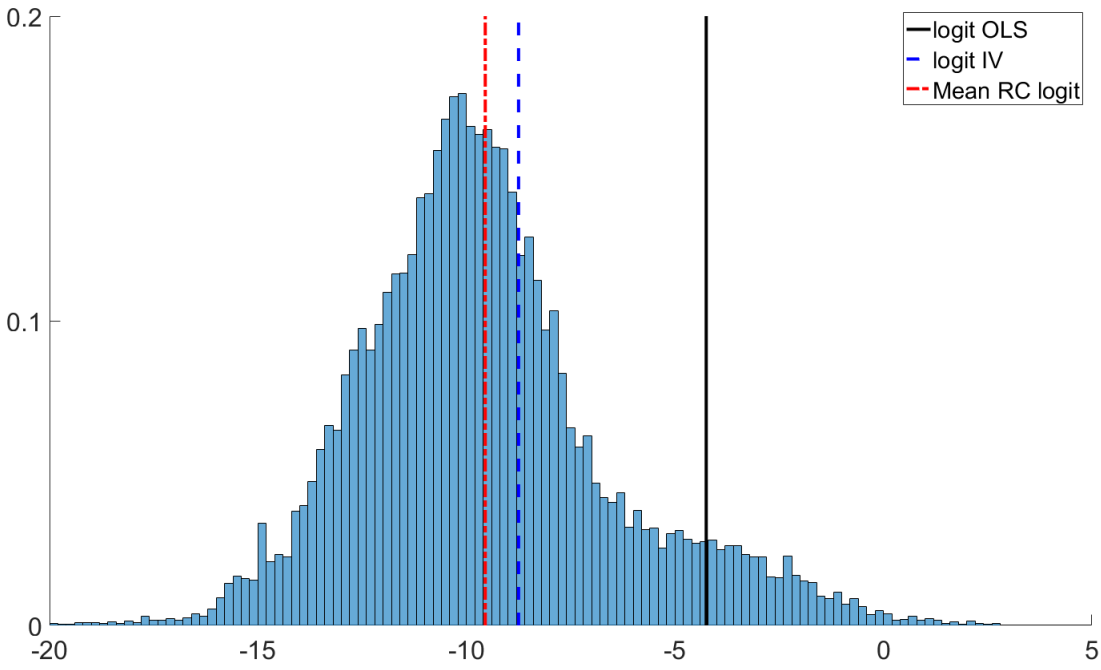


FIGURE 2. Distribution of Price Coefficient

While the price coefficients across these discrete choice models can be compared, it is not possible to directly compare them to a similar set of coefficients from the multistage AIDS specifications as there is no one price coefficient to compare them to as each equation within each segment has

several price coefficients (see the terms $\gamma_{jj'_{fm}}, g_{ff'_m}$ etc. in equations (14) which all correspond to price coefficients). Nonetheless, we provide below coefficients on the price and segment expenditure variables from one specific middle level segment – level 2, the three forms within molecule MPH – as estimates from this segment best highlight the important differences across discrete choice models when compared to product space models. To be clear, the selected segment is MPH and the set of three share equations u_1, u_2 and u_3 are for shares of MPH-IR, MPH-ER and MPH-XR (i.e., the 4-hr, 8-hr and 12-hr drugs respectively) within the methylphenidate class. Together they form a system of equations that can be estimated jointly via SUR or 3SLS, and with or without any restrictions implied by micro theory (adding up, homogeneity and symmetry).

TABLE 3. Selected Regression Coefficients for MPH Segment

	(i)			(ii)			(iii)		
	u_1	u_2	u_3	u_1	u_2	u_3	u_1	u_2	u_3
$\ln(y/P)$	-0.029	-0.006	0.036	-0.016	0.006	0.010	-0.024	-0.001	0.025
	(.002)	(.002)	(.003)	(.005)	(.004)	(.007)	(.003)	(.002)	(.004)
$\ln p_1$	-0.007	0.01	-0.003	-0.182	-0.008	0.191	-0.365	-0.170	0.536
	(.018)	(.017)	(.023)	(.170)	(.154)	(.251)	(.135)	(.048)	(.118)
$\ln p_2$	-0.044	-0.061	0.105	-0.032	-0.136	0.168	-0.170	-0.246	0.416
	(.022)	(.021)	(.028)	(.085)	(.077)	(.126)	(.048)	(.040)	(.057)
$\ln p_3$	0.078	0.032	-0.11	0.794	0.610	-1.40	0.536	0.416	-0.952
	(.019)	(.018)	(.024)	(.173)	(.157)	(.255)	(.118)	(.057)	(.126)

Note: u_1, u_2, u_3 are share equations for MPH-IR, MPH-ER and MPH-XR respectively and p_1, p_2, p_3 are prices of these three forms. Standard errors are in parenthesis. Specification (i) is SUR and does not impose homogeneity or symmetry restrictions, (ii) is GMM/IV and does not impose homogeneity or symmetry restriction and (iii) is GMM/IV and imposes both sets of restrictions. All regressions include state dummies, variables for time and time square plus additional variables at area level.

Results are summarized in Table 3. The first set of estimates for the three forms are under the columns marked as (i) and are from SUR estimation and without imposing any cross-equation restrictions. The second group of estimates (under ii) are estimated via 3SLS estimator but without imposing any restrictions. The final set of estimates (labeled iii) additionally impose homogeneity and symmetry restrictions. The restriction test has a chi-square value of 8.36 with three degrees of freedom for a p-value of 0.039. Thus, the restrictions cannot be rejected at the 1% significance level, and hence we take this as the preferred specification since without it does not make much theoretical sense.

An important result in this table is the negative coefficient on price of MPH-ER (p_2) in the equation for MPH-IR and the coefficient on price of MPH-IR (p_1) in the the equation for MPH-ER (for instance, in specification (iii) these are -0.170 and significant) implying that these forms are

complements rather than substitutes. By comparison, both forms are substitutes for the third form MPH-XR, and in turn that form is a substitute for each of these first two forms (positive coefficients of 0.536 and 0.416 respectively). Since the dependent variables in each of these equations are relative shares, and not log quantities, the coefficients cannot be interpreted as elasticities. Table (4) gives the conditional (level 2) Hicksian elasticities for this segment, which shows the complementarity between IR and ER drugs, and are calculated from the estimates from 3SLS with restrictions (estimates are significantly different from zero).

TABLE 4. MPH Hicksian Elasticities

	(A)	(B)	(C)
(A) MPH-IR	-1.66	-0.33	1.99
(B) MPH-ER	-0.61	-2.16	2.77
(C) MPH-XR	1.38	1.03	-2.41

Our conjecture is that the complementarities arise because the 4-hr and 8-hr drugs can be used in combination within a day to provide a day long coverage: a child can be given an 8-hr medication to last them through the school and then be given an additional dosage in early evening to provide coverage for the rest of the day. Taken together the two form a substitute group for the more expensive day long 12-hr drug (i.e., Concerta). This general sign pattern – complementarities between MPH-IR and MPH-ER and substitution with MPH-XR – is robust across a number of alternative specifications (and is reported elsewhere, see Bokhari and Fournier (2013)). We compare the substitution patterns across all drugs in the next section.

5.2. **Substitution Patterns.** Tables (5) provides average values of own- and cross-price elasticities of the first 11 drugs based on estimates for the discrete choice models (3) and (4), and where the average is over the 778 markets. Elasticities from the multinomial logit are given in the top part of the table. Since for the logit model cross-price elasticity for a given drug with respect to the price of any other drug is the same across all drugs (due to the IIA property), for this model we do not list the full matrix, but rather just two rows showing the own- and cross-price elasticity for each drug. The lower part of the table lists elasticities from the random coefficients logit model, including all the cross-price elasticities as they are not the same with respect to other drugs.

In the class of discrete choice logit models, the price coefficient α increased across the four specifications from -2.07 in specification (1) to -9.55 in specification (4). Consequently, the magnitudes of own- and cross-price elasticities also increase across these models. Further, model (4) overcomes the IIA limitation, leading to generally higher cross-price elasticities within a molecule-form class when compared to the simple logit models with homogenous taste parameters. The substitution patterns in (4) seem reasonable to the extent that the degree of substitutability varies by molecule and form.

TABLE 5. Elasticities from Discrete Choice Models (3 and 4)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Logit (Model 3)											
Own	-4.24	-3.00	-2.95	-5.57	-4.81	-4.80	-4.35	-8.14	-2.16	-2.27	-4.04
Cross	0.053	0.055	0.139	0.036	0.037	0.013	0.054	0.412	0.315	0.163	0.315
Random Coefficients Logit (Model 4)											
MPH-IR (4hr drugs)											
(1)	-4.50	0.074	0.188	0.040	0.044	0.015	0.061	0.406	0.356	0.197	0.240
(2)	0.071	-3.41	0.193	0.041	0.047	0.016	0.065	0.420	0.364	0.201	0.305
(3)	0.072	0.077	-3.34	0.041	0.047	0.016	0.065	0.420	0.365	0.201	0.302
MPH-ER (8hr drugs)											
(4)	0.055	0.058	0.146	-5.50	0.040	0.014	0.054	0.397	0.300	0.172	0.288
(5)	0.056	0.062	0.155	0.038	-5.04	0.015	0.058	0.418	0.316	0.179	0.328
(6)	0.057	0.062	0.157	0.038	0.045	-5.10	0.058	0.415	0.318	0.180	0.327
(7)	0.060	0.067	0.170	0.039	0.046	0.016	-4.89	0.414	0.335	0.182	0.348
MPH-XR (12hr drugs)											
(8)	0.043	0.048	0.118	0.034	0.038	0.013	0.047	-7.18	0.269	0.157	0.308
MAS-IR (4hr drugs)											
(9)	0.060	0.064	0.162	0.036	0.039	0.014	0.058	0.385	-2.20	0.186	0.287
(10)	0.031	0.044	0.093	0.045	0.065	0.016	0.029	0.570	0.160	-2.35	0.446
MAS-XR (12hr drugs)											
(11)	0.027	0.042	0.100	0.031	0.051	0.016	0.043	0.538	0.284	0.158	-4.40

Note: (1) Ritalin; (2) Methylin; (3) MPH-IR (Generics); (4) Ritalin SR/LA; (5) Metadate ER/CD; (6) Methylin ER; (7) MPH-ER (Generics); (8) Concerta; (9) Adderall; (10) MAS-IR (Generics); (11) Adderall XR.

By contrast, Table (6) lists the the substitution patterns corresponding to the multilevel AIDS model described earlier. Observe that across the two types of models (random coefficients vs nested AIDS), the own-price elasticity magnitudes decrease but the cross-price elasticity magnitudes generally increase quite dramatically. Moreover, several of the cross-price elasticities, especially those for drugs in the MPH-IR and MPH-ER segments are negative. The source of these negative cross-price elasticities is due to the complementarities in the MPH segment noted earlier.

5.3. **Merger Simulations.** To understand how important the differences in the estimated substitution patterns can be in the context of competition economics, we compare predicted percentage change in prices of all drugs under three hypothetical mergers but based on elasticity estimates from the last three models, i.e. simple logit, random coefficients logit, and nested AIDS models.

In the first hypothetical merger case, Novartis merges with Mallinckrodt. Novartis produces a 4-hr and an 8-hr variant of Ritalin, which is an established and well known brand, while Mallinckrodt competes with Novartis in both of these subsegments. Nonetheless, the market share by value of each of these firms in 2003 is small (Novartis is 3.57% and Mallinckrodt is 1.9%) and would not necessarily draw major scrutiny from competition authorities. For the second hypothetical merger,

TABLE 6. Elasticities from Nested AIDS Model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
MPH-IR (4hr drugs)											
(1)	-2.35	0.02	0.17	-0.14	-0.15	-0.05	-0.22	1.54	0.19	0.25	0.31
(2)	0.03	-0.64	-1.48	-0.14	-0.14	-0.05	-0.21	1.49	0.19	0.25	0.30
(3)	0.10	-0.59	-1.48	-0.13	-0.13	-0.05	-0.20	1.41	0.17	0.23	0.28
MPH-ER (8hr drugs)											
(4)	-0.21	-0.24	-0.57	-3.46	-0.05	0.91	0.20	2.21	0.20	0.26	0.31
(5)	-0.24	-0.28	-0.67	-0.16	-2.38	-0.16	-0.14	2.61	0.23	0.31	0.37
(6)	-0.19	-0.22	-0.52	2.42	-0.26	-3.30	-1.04	2.01	0.18	0.24	0.28
(7)	-0.18	-0.21	-0.50	0.21	0.10	-0.26	-2.14	1.93	0.17	0.23	0.27
MPH-OROS (12hr drugs)											
(8)	0.19	0.22	0.54	0.20	0.21	0.08	0.31	-3.02	0.21	0.27	0.33
MAS-IR (4hr drugs)											
(9)	0.05	0.06	0.13	0.03	0.03	0.01	0.05	0.34	-2.52	1.06	0.10
(10)	0.05	0.06	0.14	0.03	0.04	0.01	0.05	0.36	0.75	-2.32	0.11
MAS-XR (12hr drugs)											
(11)	0.06	0.06	0.15	0.04	0.04	0.01	0.06	0.40	0.02	0.03	-1.65

Note: (1) Ritalin; (2) Methylin; (3) MPH-IR (Generics); (4) Ritalin SR/LA; (5) Metadate ER/CD; (6) Methylin ER; (7) MPH-ER (Generics); (8) Concerta; (9) Adderall; (10) MAS-IR (Generics); (11) Adderall XR.

we consider a medium size merger, where Novartis with its small market share merges with Alza with 26.10% share for its product Concerta (the drug was developed by Alza and was distributed by Ortho-McNeil, but for simplicity henceforth we just refer to the firm as Alza). At the time, Alza was the only company selling the 12-hr variant of the MPH molecule but it had no products in the competing segments of 4-hr and 8-hr drugs within this molecule. Hence we consider a merger between Alza in the 12-hr segment in MPH with Novartis providing 4-hr and 8-hr drugs. Such a merger would likely draw some attention from competition authorities and invoke further investigation. The third and last hypothetical merger we consider is between Alza and Shire. The market share of Shire through its three products in competing molecule class of MAS and DEX was 26.89%. Also, Shire was the only other firm producing a day-long drug, Adderall XR, albeit for a different molecule than Concerta. A proposed merger between two largest firms would definitely be investigated by the competition authorities.

Our merger simulations are based on a Nash-Bertrand price competition model in the context of multiproduct firms, where based on an estimated demand model, we first back out marginal costs via an equilibrium price equation, and then use these costs to compute new prices under joint profit maximization of co-owned products of the merging firms (see Nevo, 1998). To be clear, let there be J related products where the unconditional demand for product j is given by

$$Q_j = D_j(p_1, \ldots, p_J, Z) \tag{16}$$

and where Z represents exogenous demand shift variables. If there are L firms, and the lth firm produces a subset $\mathfrak{L}_l$ of the products, then it maximizes over the sum of profits associated with each products as

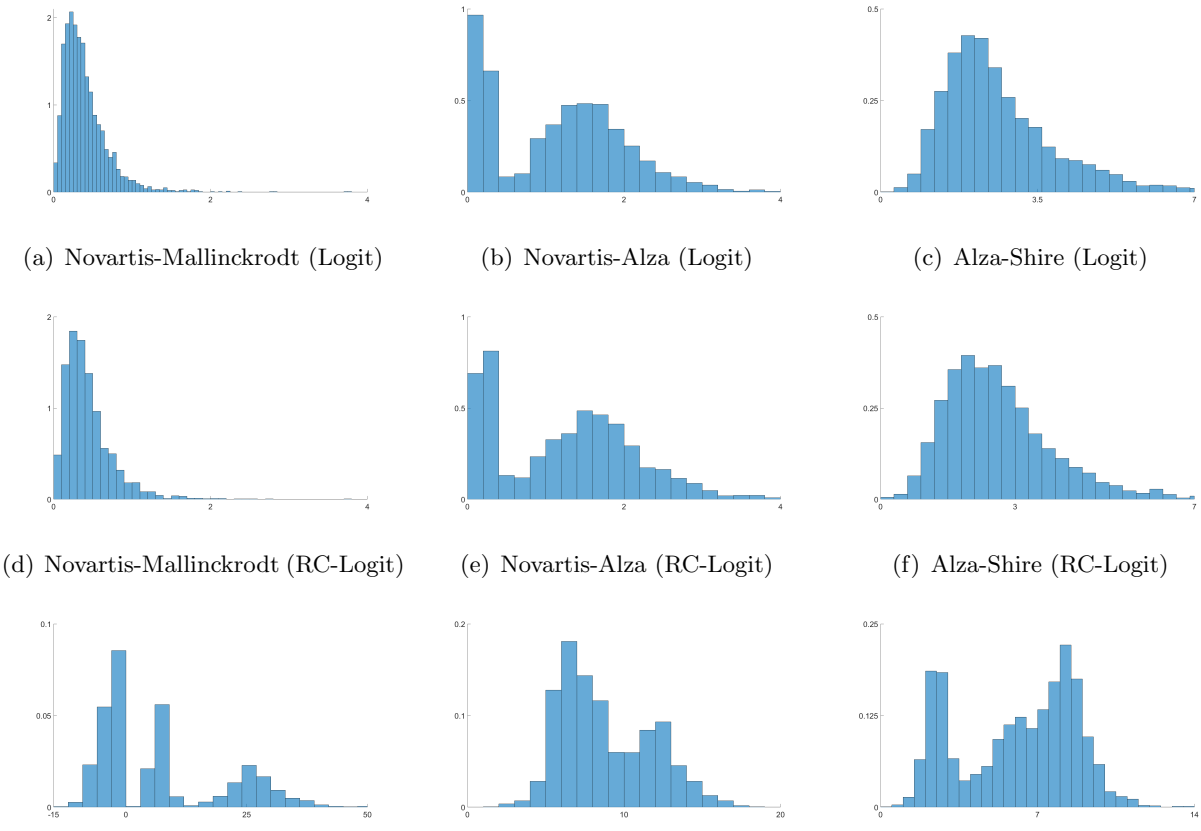
$$\Pi_l = \sum_{r \in \mathfrak{L}_l} (p_r - c_r) D_r(p_1, \ldots, p_J, Z)$$

where c_r is the constant marginal cost. Under Nash-Bertrand price competition, price p_j of any product j produced by firm l satisfies the first order conditions

$$Q_j + \sum_{r \in \mathfrak{L}_l} (p_r - c_r) \frac{\partial D_r(p_1, \ldots, p_J, Z)}{\partial p_j} = 0. \tag{17}$$

If we let Θ be a 1/0 matrix with ones in the leading diagonal and in locations where a firm jointly produces products r and j and define Ω such that $\Omega_{jr} = -\Theta_{jr} \frac{\partial D_r(p_1, \ldots, p_J, Z)}{\partial p_j}$, then the first order conditions imply a price equation (in matrix notation) of the form

$$p = c + \Omega^{-1} D(p_1, \ldots, p_J, Z). \tag{18}$$



(a) Novartis-Mallinckrodt (Logit) (b) Novartis-Alza (Logit) (c) Alza-Shire (Logit)

(d) Novartis-Mallinckrodt (RC-Logit) (e) Novartis-Alza (RC-Logit) (f) Alza-Shire (RC-Logit)

(g) Novartis-Mallinckrodt (AIDS) (h) Novartis-Alza (AIDS) (i) Alza-Shire (AIDS)

FIGURE 3. Predicted percentage price change

For each of the three models, we proceed as follows. Using demand parameters and data from 2003, and the actual ownership matrix Θ, we back out the marginal cost for each product in each market using equation (18) above. This calculation involves solving 17 linear equations in each market. Next, change the ownership matrix to match one of the three hypothetical mergers given above. Based on the estimates of marginal costs and the new ownership matrix, solve for new equilibrium prices in each market per equation (18) again. Since the markup term $\Omega^{-1}D(p_1, \ldots, p_J, Z)$ is also a function of the price vector, this requires solving a non-linear system of 17 equations per market.[15]

TABLE 7. Three Hypothetical Mergers

				% Mean† Change in Price								
		2003 Data		I. Novartis-Mallinckrodt			II. Novartis-Alza			III. Alza-Shire		
Firm	Product	Price	Share	Logit	RC-Logit	AIDS	Logit	RC-Logit	AIDS	Logit	RC-Logit	AIDS
Novartis	(1) Ritalin	52.41	0.88%	0.31	0.36	-0.73	1.72	1.85	8.11	0.01	-0.02	2.02
Mallinckrodt	(2) Methylin	31.81	1.27%	0.64	0.68	-5.69	0.00	0.00	1.52	0.01	-0.02	2.29
19 firms	(3) Generic MPH-IR	33.06	2.65%	0.00	0.00	0.50	0.00	0.00	2.20	0.02	-0.01	3.10
Novartis	(4) Ritalin SR/LA	59.20	2.69%	0.27	0.30	6.87	1.52	1.65	12.29	0.01	-0.02	2.33
Celltech	(5) Metadate CD/ER	57.18	2.61%	0.00	0.00	0.52	0.00	0.00	2.61	0.01	-0.02	3.13
Mallinckrodt	(6) Methylin ER	52.29	0.63%	0.38	0.40	26.99	0.00	0.00	5.98	0.01	-0.02	2.69
15 firms	(7) Generic MPH-ER	48.12	0.81%	0.00	0.00	0.47	0.00	0.00	2.39	0.00	-0.03	2.86
Alza	(8) Concerta	96.45	26.10%	0.00	0.00	0.45	0.20	0.24	6.03	2.05	2.91	5.45
Shire	(9) Adderall	31.87	2.94%	0.00	-0.08	0.09	0.01	-0.08	0.69	2.91	3.05	2.52
3 firms	(10) Generic MAS-IR	26.52	7.61%	0.00	0.00	0.14	0.01	0.00	1.02	0.07	0.02	1.49
Shire	(11) Adderall XR	52.58	23.78%	0.00	0.00	0.19	0.01	0.01	1.41	1.74	1.74	8.17
GSK	(12) Dexedrine	34.94	0.19%	0.00	0.00	0.02	0.00	0.00	0.15	0.00	-0.02	0.61
Shire	(13) Dextrostat	23.25	0.17%	0.00	0.01	0.24	0.02	0.02	1.73	3.94	3.33	8.22
4 firms	(14) Generic DEX-IR	24.62	0.40%	0.00	0.00	0.05	0.00	0.00	0.38	0.00	-0.03	1.59
GSK	(15) Dexedrine SR	40.25	0.70%	0.00	0.00	0.05	0.00	0.00	0.40	0.00	-0.02	1.65
2 firms	(16) Generic DEX-ER	35.14	1.10%	0.00	0.00	0.06	0.00	0.00	0.43	0.01	-0.02	1.77
11 firms	(17) All Other drugs	85.18	25.48%	0.00	0.00	0.00	0.00	0.00	0.00	0.02	-0.01	0.00
	Average (Weighted)			0.02	0.02	0.49	0.11	0.12	2.62	1.03	1.26	3.91

†Mean is over all counties in 2003. For 95 percentile range see appendix.

[15] If the demand system is linear, these equations also become linear as then the slope matrix $\frac{\partial D_r(\,)}{\partial p_j}$ is not a function of prices. For the multilevel AIDS model, we took advantage of this simplification by assuming that the overall unconditional demand system is locally linear. Then by using the mean prices and shares/quantities in 2003, we converted the unconditional elasticity matrix from multilevel AIDS model given in Table 6 to obtain estimates of intercepts and slopes for the linear demand system. This simplification primarily allows us to avoid predicting quantities and relative and absolute shares (at incremental price vector) using large number of demand equations associated with the multilevel AIDS system. We further verified that this simplification does not lead to any major differences in the final results by subjecting the other two models to the same linear simplification, i.e., by backing intercepts and slopes for the unconditional demand system by staring out with the elasticity matrix from the other two models. The merger simulation results for the logit and random coefficients logit models were very similar to when we use the correct non-linear system implied by the equation above or via this linear unconditional demand. There were some additional (minor) changes made to the multilevel AIDS elasticity matrix for the merger analysis. They do not change the main conclusions of this section and are described in the appendix.

The results of hypothetical mergers, i.e., predicted percentage change in price for drugs associated with merging parties are summarized as 3×3 plots in Figure 3. The first column in the figure shows the percentage price changes for four drugs owned by Novartis and Mallinckrodt (i.e., price changes for drugs 1,2,4, and 6). The three rows summarize the distribution of price changes over the 778 counties used in the simulations and across the three demand models. Similarly, the second and third columns of the figure show distribution of prices changes for drugs associated with Novartis-Alza (drugs 1,4, and 8) and Alza-Shire (drugs 8,9,11 and 13) mergers. The mean price changes for all drugs, not just the drugs associated with the merging companies are summarized in Table 7 (and the 95 percentile interval for each drug and merger is given in the appendix in Table A-1). Under each merger, the logit model predicts very small changes in the price of the drugs of the merging companies, and almost none in the price of all other products. By comparison, the random coefficients model predicts slightly larger changes in the prices of the drugs of the merging companies. In terms of overall average price changes (weighted by share of sales pre-merger) the percentage change across the two models is very similar in this application. However, the predicted price changes are significantly much larger for the multilevel AIDS model and the distributions of changes are quite different compared to the other two models, not only for the drugs of the merging companies but also of other drugs with share weighted average ranging from 0.49 for the first (small) merger to 3.91 for the third (large) merger. The exception is the first merger involving firms with small market shares, but one that also involves drugs that are complementary segments. In this case, the AIDS model predicts a small price decrease for some of the four hour drugs. More generally though, under the third demand model the predicted prices changes of the merging parties as well as of the competitors are both quite large.

5.4. **Limitations.** While the results above highlight comparisons across different demand specifications – which is the primary purpose of this paper – some limitations relating to the aggregate nature of the data and the associated measurement of price in the presence of health insurance are worth noting prior to using reported elasticities in other works.[16]

The price of a drug is computed using the ratio of revenue to quantity sold in a given county. Revenue is the sum of monies received by retail pharmacies from all parties, and is inclusive of co-pay from the patient, as well as payment from an insurance company or any third parties such as Medicaid paying on behalf of the patient. In this respect it approximates the average transaction price for the drug in a county. However, due to the presence of insurance, the actual price varies across different payers (patient plus insurer) within a market. This is because Pharmacy Benefit Managers (PBMs) make extensive use of multi-tiered formularies to control utilization, where brands placed in tier-one have a smaller co-payment relative to those in higher tiers. In exchange for a discount from a given manufacturer, a PBM can place a specific brand on tier-one with low co-pay,

[16]We are in debt to an anonymous referee for alerting us to this issue.

while the competing brands may be on higher tiers and consequently require higher co-payments (see Berndt and Newhouse, 2010). Similarly, a different insurance company in the same market may have a different brand on tier-one. In turn this implies that the actual price of a product j paid by payer n (i.e., patient-insurer combination in our context) differs from the average price in an area t due to the presence of multi-tiered formularies in the US health insurance market.[17] To be clear, because of insurance formularies different consumers face different relative prices but in the aggregate data it is assumed that all consumers within a market face the same relative prices.

An additional point worth acknowledging is that we have abstracted away from the agency problem where physicians influence the choice of a given brand and are targeted by pharmaceuticals via marketing, while the patient and insurer collectively pay for it. Even if physicians are assumed as perfect agents, given that there are two payers, one can ask what does the area under the demand curve represent? One way to think about the area under the demand curve is the sum of consumer (patient) surplus plus that of the insurance company, and without explicit data on co-payments for each sale it cannot be split into its two parts, i.e. how much of it is appropriated by insurers relative to consumers. Nonetheless, the change in the sum of these two surpluses can still provide a useful metric when considering the impact of a price change due to a policy or a merger on the joint patient/insurer consumer (see Branstetter et al., 2016, p.886, for a similar interpretation).

To address these issues is beyond the scope of the current paper. Our goal was to use conventional demand estimation techniques used in the empirical literature (e.g., Hausman and Leonard, 2005, Ivaldi and Verboven, 2005) to highlight the importance of functional form in estimating demand, and in simulating the price effects of mergers. In future research it would also be valuable to explore the aggregation issues, and estimate demand using individual level data to examine the extent of aggregation bias in estimating demand for pharmaceuticals.

6. Conclusion

For the ADHD data used in this paper, we believe that a flexible demand system such as AIDS is more appropriate specification than a discrete choice model. As documented earlier, the data requirements for estimating product space models are substantial and practical limitations may favor the use of discrete choice models. However, as long as the analysis is for a handful of products, i.e., drugs within a narrowly defined therapeutic area, the problem of large number of parameters associated with AIDS models may not be insurmountable. On the other hand, a discrete choice model cannot, without some additional information, account for underlying complementarities which may be important in some industries such as the pharmaceuticals (see for instance Gentzkow (2007)).

[17] The average price problem we discuss here is common to all papers that use similarly aggregated data, but where price varies by individuals due to insurance or other discounts to the payers. Examples from pharmaceuticals from the US with aggregate data include (among others) Ellison et al. (1997), Cleanthous (2002), Branstetter et al. (2016), Minjae et al. (2017).

In the examples used in this paper, discrete choice models give very different, and typically smaller, predicted changes in prices for merging drugs as well as for the competitors' drugs than the AIDS models. They can also miss a *decrease* in prices that may follow in case of complementarities. These differences across models are driven by both the magnitude and sign of cross-price elasticities.

Our results (for the magnitude at least) are consistent with the one reported in Weinberg and Hosken (2013) who also find that merger predictions were larger for AIDS demand compared to the logit models. Insights from our work also provide a potential explanation for merger results reported in Björnerstedt and Verboven (2016). They use nested logit models to predict post-merger prices in the Swedish analgesics market, and compare them to observed prices post an actual merger. The authors report that while overall the simulations were quite successful in terms of closeness between predicted and true price changes, nonetheless the simulations slightly under-predicted the price increase for the merging firms, and significantly under-predicted the price increase by the non-merging competitors. We also find that predicted price increase for the hypothetical merger for some of the non-merging competitors is considerably larger under AIDS than for the other two models. Thus we conjecture that the differences between true and predicted price changes observed in the Swedish analgesics market may be similar to the one we observe here between discrete choice and AIDS models, and may be related to the difference in magnitude of cross-price elasticities across these models.

On the flip side however, the magnitude of the cross-price elasticities in AIDS models can be too 'large', in the sense that they may lead to the failure of second order conditions for profit maximization not holding in merger simulations. This can happen for instance when the sum of the cross-price elasticities is larger than the own-price elasticities, and the ownership matrix also mostly consists of ones in the off-diagonals (this potential problem does not occur in logit or random coefficients models as severely). While predicted price changes are unlikely to be the only factor a competition authority would consider when deliberating a merger, nonetheless, these examples show that multilevel AIDS model can give very different estimates from discrete choice models. The results of this paper are meant to serve as a cautionary tale when using merger simulations, as choice of the demand model matters, and not an overall indictment against them or against any one type of demand models.

References

Abdul-Ghani, M.A., C. Puckett, C. Triplitt, D. Maggs, J. Adams, E. Cersosimo, and R.A. DeFronzo, "Initial combination therapy with metformin, pioglitazone and exenatide is more effective than sequential add-on therapy in subjects with new-onset diabetes. Results from the efficacy and durability of initial combination therapy for Type 2 diabetes (EDICT): a randomized trial," *Diabetes, Obesity and Metabolism*, March 2015, *17* (3), 268–275.

Armstrong, Mark, "A more general theory of commodity bundling," *Journal of Economic Theory*, March 2013, *148* (2), 448–472.

Barten, A. P., "Consumer demand functions under conditions of almost additive preferences," *Econometrica*, January-April 1964, *32* (1/2), 1–38.

Berndt, Ernst R. and Joseph P. Newhouse, "Pricing and reimbursement in U.S. pharmaceutical markets," NBER Working Paper w16297, National Bureau of Economic Research August 2010.

Berry, Steven, J. Levinsohn, and Ariel Pakes, "Automobile prices in market equilibrium," *Econometrica*, July 1995, *63* (4), 841–890.

Berry, Steven T., "Estimating discrete-choice models of product differentiation," *RAND Journal of Economics*, Summer 1994, *25* (2), 242–262.

_____ **and Philip A. Haile**, "Identification in Differentiated Products Markets Using Market Level Data," *Econometrica*, September 2014, *82* (5), 1749–1797.

Björnerstedt, Jonas and Frank Verboven, "Does merger simulation work? Evidence from the Swedish analgesics market," *American Economic Journal: Applied Economics*, July 2016, *8* (3), 125–164.

Bokhari, Farasat A. S. and Gary M. Fournier, "Entry in the ADHD drugs market: Welfare impact of generics and me-toos," *Journal of Industrial Economics*, June 2013, *61* (2), 340–393.

_____ **and Helen Schneider**, "School accountability laws and consumption of psychostimulants," *Journal of Health Economics*, March 2011, *30* (2), 355–372.

Branstetter, Lee G., Chirantan Chatterjee, and Matthew Higgins, "Regulation and welfare: Evidence from paragraph IV generic entry in the pharmaceutical industry," *RAND Journal of Economics*, Winter 2016, *47* (4), 857–890.

Bresnahan, Timothy F., "Comment to: Valuation of New Goods under Perfect and Imperfect Competition," in Robert James Gordon Timothy F. Bresnahan, ed., *The Economics of New Goods*, Vol. 58 of *Studies in Income and Wealth*, National Bureau of Economic Research, 1997, pp. 237–248.

Budzinski, Oliver and Isabel Ruhmer, "Merger simulation in competition policy: A survey," *Journal of Competition Law and Economics*, June 2010, *6* (2), 277–319.

Cardell, N. Scott, "Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity," *Econometric Theory*, April 1997, *13* (2), 185–213.

Chaudhuri, Shubham, Pinelopi K. Goldberg, and Panle Jia, "Estimating the effects of global patent protection in pharmaceuticals: a case study of quinolones in India," *American Economic Review*, December 2006, *96* (5), 1477–1514.

Christensen, Laurits R., Dale W. Jorgenson, and Lawrence J. Lau, "Transcendental logarithmic utility functions," *American Economic Review*, June 1975, *65* (3), 367–383.

Cleanthous, Paris, "Patient welfare implications of innovation in the U.S. antidepressant Market," mimeo 2002. https://tinyurl.com/yaqabezm.

Conner, Daniel F., "Stimulants," in Russell A. Barkley, ed., *Attention-deficit hyperactvity disorder: A handbook for diagnosis and treatment*, 3 ed., Guilford Press, 2006, chapter 17, pp. 608–647.

Crooke, Philip, Luke Froeb, Steven Tschantz, and Gregory J. Werden, "Effects of assumed demand form on simulated postmerger equilibria," *Review of Industrial Organization*, November 1999, *15* (3), 205–217.

Deaton, Angus and John Muellbauer, "An almost ideal demand system," *American Economic Review*, June 1980, *70* (3), 312–326.

Dubé, Jean-Pierre, Jeremy T. Fox, and Che-Lin Su, "Improving the numerical performance

of static and dynamic aggregate discrete choice random coefficients demand estimation," *Econometrica*, September 2012, *80* (5), 2231–2267.

Duso, Tomaso, Annika Herr, and Moritz Suppliet, "The welfare impact of parallel imports: A structural approach applied to the German market for oral anti-diabetics," *Health Economics*, September 2014, *23* (9), 1036–1057.

Ellison, Sara Fisher, Iain Cockburn, Zvi Griliches, and Jerry A. Hausman, "Characteristics of demand for pharmaceutical products: an examination of four cephalosporins," *RAND Journal of Economics*, Autumn 1997, *28* (3), 426–446.

Fox, Jeremy T. and Natalia Lazzati, "A note on identification of discrete choice models for bundles and binary games," 2016. Cemmap working paper CWP04/13.

Gentzkow, Matthew, "Valuing new goods in a model with complementarity: Online newspapers," *American Economic Review*, June 2007, *97* (3), 713–744.

Grigolon, Laura and Frank Verboven, "Nested logit or random coefficients logit? A comparison of alternative discrete choice models of product differentiation," *The Review of Economics and Statistics*, December 2014, *96* (5), 916–935.

Hausman, Jerry A. and Gregory K. Leonard, "The competitive effects of a new product introduction: A case study," *Journal of Industrial Economics*, September 2002, *L* (3), 237–263.

_____ **and** _____, "Competitive analysis using a flexible demand specification," *Journal of Competition Law and Economics*, 2005, *1* (2), 279–301.

_____**, G. Leonard, and J.D. Zona**, "Competitive analysis with differentiated products," *Annales d'Economie et de Statistique*, 1994, *34*, 159–180.

Hay, Alastair D., Céire Costelloe, Niamh M. Redmond, Alan A. Montgomery, Margaret Fletcher, Sandra Hollinghurs, and Tim J. Peters, "Paracetamol plus ibuprofen for the treatment of fever in children (PITCH): randomised controlled trial," *British Medical Journal*, September 2008, *337*, a1302.

Huang, Dongling, Christian Rojas, and Frank Bass, "What happens when demand is estimated with a misspecified model?," *The Journal of Industrial Economics*, December 2008, *56* (4), 809–839.

Ivaldi, Marc and Frank Verboven, "Quantifying the effects from horizontal mergers in European competition policy," *International Journal of Industrial Organization*, December 2005, *23* (9-10), 669–691.

Knittel, Christopher R. and Konstantinos Metaxoglou, "Estimation of random-coefficient demand models: Two empiricists' perspective," *The Review of Economics and Statistics*, March 2014, *96* (1), 34–59.

Lancaster, Kelvin J., "A New Approach to Consumer Theory," *Journal of Political Economy*, April 1966, *74* (2), 132–157.

Mayoral, Clara E., Ronald V. Marino, Warren Rosenfeld, and Joseph Greensher, "Alternating antipyretics: is this an alternative?," *Pediatrics*, May 2000, *105* (5), 1009–1012.

McFadden, Daniel, "Conditional logit analysis of qualitative choice behavior," in Paul Zarembka, ed., *Frontiers in econometrics*, Academic Press, 1973, chapter 4, pp. 105–142.

Miller, Nathan H., Marc Remer, Conor Ryan, and Gloria Sheu, "Pass-through and the prediction of merger price effects," *The Journal of Industrial Economics*, December 2016, *64* (4), 683–709.

_____**,** _____**,** _____**, and** _____, "Upward pricing pressure as a predictor of merger price effects," *International Journal of Industrial Organization*, May 2017, *52*, 216–247.

Minjae, Song, Sean Nicholson, and Claudio Lucarelli, "Mergers with interfirm bundling: a case of pharmaceutical cocktails," *The RAND Journal of Economics*, Fall 2017, *48* (3), 810–834. Previously ciculated as "Previously circulated as "Bundling Among Rivals: A Case of Pharmaceutical Cocktails" as NBER wp.

Moon, Hyungsik Roger, Matthew Shum, and Martin Weidner, "Estimation of random coefficients logit demand models with interactive fixed effects," Cemmap Working Paper CWP20/14, Institute for Fiscal Studies 2014.

Nevo, Aviv, "Identification of the oligopoly solution concept in a differentiated-products industry," *Economics Letters*, 1998, *59*, 391–395.

_____, "Mergers with differentiated products: The case of the ready-to-eat cereal industry," *RAND Journal of Economics*, Autumn 2000, *31* (3), 395–421.
_____, "A practitioner's guide to estimation of random-coefficients logit models of demand," *Journal of Economics and Management Strategy*, Winter 2000, *9* (4), 513–548.
_____, "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, March 2001, *69* (2), 307–342.
_____**, Daniel L. Rubinfeld, and Mark McCabe**, "Academic journal pricing and the demand of libraries," *The American Economic Review, Papers and Proceedings*, May 2005, *95* (2), 447–452.
Peters, Craig, "Evaluating the performance of merger simulation: Evidence from the U.S. airline industry," *Journal of Law and Economics*, October 2006, *49* (2), 627–649.
Petrin, Amil, "Quantifying the benefits of new products: The case of the minivan," *Journal of Political Economy*, August 2002, *110* (4), 705–729.
Pinkse, Joris and Margaret E. Slade, "Mergers, brand competition, and the price of a pint," *European Economic Review*, June 2004, *48* (3), 617–643.
_____**,** _____**, and Craig Brett**, "Spatial price competition: A semiparametric approach," *Econometrica*, May 2002, *70* (3), 1111–1153.
Reynaert, Mathias and Frank Verboven, "Improving the performance of random coefficients demand models: The role of optimal instruments," *Journal of Econometrics*, March 2014, *179* (1), 83–98.
Rojas, Christian, "Price competition in U.S. brewing," *Journal of Industrial Economics*, March 2008, *56* (1), 1–31.
Skrainka, Benjamin S. and Kenneth L. Judd, "High performance quadrature rules: How numerical integration affects a popular model of product differentiation," Cemmap working paper CWP03/11, Institute for Fiscal Studies 2011.
Stone, Richard, "Linear Expenditure Systems and demand analysis: An application to the pattern of British demand," *The Economic Journal*, September 1954, *64* (255), 511–527.
Sudhir, K., "Structural analysis of manufacturer pricing in the presence of a strategic retailer," *Marketing Science*, Summer 2001, *20* (3), 244–264.
Theil, H., "The information approach to demand analysis," *Econometrica*, January 1965, *33* (1), 67–87.
Verboven, Frank, "International price discrimination in the European car market," *RAND Journal of Economics*, Summr 1996, *27* (2), 240–268.
Weinberg, Matthew C., "More evidence on the performance of merger simulations," *The American Economic Review, (Papers and proceedings)*, May 2011, *101* (3), 51–55.
_____ **and Daniel Hosken**, "Evidence on the accuracy of merger simulations," *The Review of Economics and Statistics*, December 2013, *95* (5), 1584–1600.

Appendix

Simulation Details. The elasticity matrix for the multilevel AIDS was modified in three ways before using it for merger analysis. First, if the cross-elasticity estimate was not statistically significantly different from zero, then we set it to zero before the merger analysis. This did not change the results in any appreciable manner for the mergers considered here, except that price changes in the third merger were even larger with average price change being 5.10% instead of 3.91% as reported in the paper. Second, the own-price elasticity of drug (2), Methylin by Mallinckrodt, was estimated to be in the inelastic region ($\eta_{2,2} = -0.64$). Using this estimate as is, leads to negative estimates of marginal costs for this drug. Since it is otherwise very similar to generic versions of MPH-IR, and indeed itself is almost a generic in terms of brand recognition and prices, we assumed that its marginal cost would be similar to that of the other generics in this segment. Hence to overcome this difficulty, we set its own-price elasticity to be the same as that of generic MPH-IR. This modification leads to some modest differences in merger predicted prices, particularly in the first hypothetical merger case as it involves merger between this firm and Novartis. Overall average price change is only 0.16% for the first merger (instead of 0.49% as reported in the paper) if we do not make change. Third, we set the cross-price elasticity of drug (17) to zero. If we do not make this change (but continue to make the other two changes), results for average percentage change in prices are even larger than those reported in the paper for all three mergers. This change was necessary because, as documented in Bokhari and Fournier (2013), demand for this last drug is measured very imprecisely and small changes in the underlying model specification change the cross-price elasticity for this drug with the other 16 drugs substantially. Thus we set the off-diagonal in the elasticity matrix associated with drug 17 to zero so that when the Nash-Bertrand equilibrium prices are computed for the remaining 16 drugs under a merger scenario, the cross-effects with drug 17 would not impact those calculations. Consequently, the price change in drug 17 is also zero under all three scenarios.

TABLE A-1. Three Hypothetical Mergers

		% Change in Price — 95 Percentile Range								
		I. Novartis-Mallinckrodt			II. Novartis-Alza			III. Alza-Shire		
Firm	Drug #	Logit	RC-Logit	AIDS	Logit	RC-Logit	AIDS	Logit	RC-Logit	AIDS
Novartis	(1)	[.05,.85]	[.05,1.06]	[-1.03,-0.48]	[.77,3.10]	[.83,3.38]	[5.48,11.3]	[.00,.02]	[-.07,.01]	[1.05,2.84]
Mallinckrodt	(2)	[.15,1.54]	[.16,1.59]	[-10.6,-3.09]	[.00,.01]	[-.00,.01]	[0.75,2.87]	[.00,.04]	[-.07,.02]	[1.12,4.03]
19 firms	(3)	[.00,.00]	[-.00,.00]	[0.21,0.88]	[.00,.01]	[-.00,.01]	[1.16,3.49]	[.00,.04]	[-.06,.03]	[1.59,4.51]
Novartis	(4)	[.04,.79]	[.05,.86]	[3.92,10.35]	[.60,2.75]	[.68,3.02]	[8.55,16.19]	[.00,.02]	[-.10,.01]	[1.15,3.31]
Celltech	(5)	[.00,.00]	[-.01,.00]	[0.33,0.72]	[.00,.00]	[-.01,.00]	[1.47,3.46]	[.00,.02]	[-.10,.00]	[1.55,4.47]
Mallinckrodt	(6)	[.10,.82]	[.12,.90]	[16.0,40.6]	[.00,.00]	[-.01,.00]	[4.40,8.75]	[.00,.02]	[-.08,.01]	[1.31,4.15]
15 firms	(7)	[.00,.00]	[-.00,.00]	[0.30,0.74]	[.00,.00]	[-.01,.00]	[1.39,3.61]	[.00,.01]	[-.09,.00]	[1.44,4.31]
Alza	(8)	[.00,.00]	[-.01,.00]	[0.29,0.61]	[.06,.43]	[.08,.60]	[3.68,7.85]	[1.01,3.27]	[1.13,5.17]	[2.17,8.04]
Shire	(9)	[.00,.00]	[-.01,.00]	[0.06,0.13]	[.00,.03]	[-.02,.03]	[0.44,0.92]	[1.17,5.45]	[1.25,5.95]	[1.46,3.76]
3 firms	(10)	[.00,.00]	[-.01,.00]	[0.09,0.20]	[.00,.02]	[-.01,.02]	[0.67,1.36]	[.02,.16]	[-.15,.15]	[0.81,2.16]
Shire	(11)	[.00,.00]	[-.00,.00]	[0.13,0.26]	[.00,.02]	[.00,.02]	[0.94,1.77]	[.76,3.07]	[.68,3.31]	[5.89,10.8]
GSK	(12)	[.00,.00]	[-.00,.00]	[0.01,0.03]	[.00,.00]	[-.00,.00]	[0.10,0.19]	[.00,.01]	[-.07,.01]	[0.46,0.78]
Shire	(13)	[.00,.01]	[.00,.01]	[0.14,0.35]	[.00,.05]	[.00,.05]	[1.19,2.31]	[1.77,7.26]	[1.35,6.19]	[5.87,10.7]
4 firms	(14)	[.00,.00]	[-.00,.00]	[0.03,0.08]	[.00,.00]	[-.00,.01]	[0.26,0.53]	[.00,.01]	[-.09,.01]	[1.20,2.15]
GSK	(15)	[.00,.00]	[-.00,.00]	[0.03,0.08]	[.00,.00]	[-.01,.00]	[0.25,0.54]	[.00,.01]	[-.08,.01]	[1.20,2.14]
2 firms	(16)	[.00,.00]	[-.00,.00]	[0.04,0.08]	[.00,.00]	[-.01,.00]	[0.29,0.56]	[.00,.02]	[-.07,.01]	[1.28,2.26]
11 firms	(17)	[.00,.00]	[-.01,.00]	[0.00,0.00]	[.00,.01]	[-.02,.01]	[0.00,0.00]	[.01,.05]	[-.13,.04]	[0.00,0.00]

Note: (1) Ritalin; (2) Methylin; (3) MPH-IR (Generics); (4) Ritalin SR/LA; (5) Metadate ER/CD; (6) Methylin ER; (7) MPH-ER (Generics); (8) Concerta; (9) Adderall; (10) MAS-IR (Generics); (11) Adderall XR.