Empirical Industrial Organisation and Market Design

Isacco Lotti, Matteo Pozzi, Apostolos Georgiadis, Selen Saygun

December 19, 2023

Computational Assignment

1 The BLP model: a general introduction

Discrete-choice demand models allow to manage rich substitution patterns between large sets of products. The model developed by Berry, Levinsohn, and Pakes (1995) introduces controls for the endogeneity of product characteristics without altering the flexibility of these substitution patterns. BLP propose a generalized method-of-moments (GMM) estimator and a numerical algorithm to find it.

The standard random coefficients, discrete-choice model considers a set of markets, $t = 1, \dots, T$ each populated by a number H_t of consumers, everyone of them choose one of the $j = 1, \dots, J$ products available, or opt for the outside option. Each product j is described by its characteristics $(x_{j,t}, \xi_{j,t}, p_{j,t})$. Where $p_{j,t}$ is the price and $x_{j,t}$ is a vector of attributes of product j. $\xi_{j,t}$ represents a vertical characteristics that is unobserved by the researcher: it can be seen as a demand shock that is market and product specific and is common across all consumers. Consumer h in market t obtains the utility from purchasing product j

$$u_{h,j,t} = \beta_h^0 + x'_{j,t}\beta_h^x - \beta_h^p p_{j,t} + \xi_{j,t} + \epsilon_{h,j,t} \ . \tag{1}$$

The parameter vector β_h^x contains the consumer's tastes for the attributes, β_i^p reflects the marginal utility of income: consumer h's price sensitivity. The intercept β_h^0 captures the value of purchasing an inside good instead of an outside good. Coefficients are consumer specific since they are aggregates of a vertical component with an horizontal one that depends from their demographics D_h and a random drawn v_h.

$$\begin{pmatrix} \beta_h^0 \\ \beta_h^x \\ \beta_h^p \end{pmatrix} = \begin{pmatrix} \beta^0 \\ \beta^x \\ \beta^p \end{pmatrix} + \Sigma v_h + \Pi D_h \ . \tag{2}$$

For the BLP algorithm in (1) it is convenient to distinguish between the vertical component of the parameters $\theta_1 = (\beta^0, \beta^x, \beta^p)$ from the others $\theta_2 = (\Pi, \Sigma)$ in order to get:

$$u_{h,j,t} = \delta_{j,t}(x_{j,t}, p_{j,t}, \xi_{j,t}; \theta_1) + \mu_{h,j,t}(x_{j,t}, p_{j,t}, D_h, v_h, \theta_2) \ . \tag{3}$$

Where $\delta_{j,t}$ is the average utility across all the H_t consumers derived by purchasing product j in market t.

The goal of the BLP algorithm is to estimate the vector of parameters θ that minimizes the norm between the market shares predicted by the model and the actual market shares. Starting from an educated guess of δ_t, e.g. using simple logit, and from an initial value for θ_2 market shares predicted by the model can be computed as

$$\sigma_t(\delta_t, x_t, p_t; \theta_2) = \int I[u_{h,j,t} \geq u_{h,k,t} \forall k \neq j] \, dF(\epsilon_{h,t}, D_{h,t}, v_{h,t}) \ . \tag{4}$$

As for the simple logit, $\epsilon_{h,t}$ are assumed to be distributed Type I extreme value so that they can be integrated out analytically. The integrals in (4) are typically evaluated by Monte Carlo simulation with n_s draws of v_h from $F_v(v)$ and of D_i from $F_D(D)$. In BLP it is proved that the function σ is invertible. Thus, given a guess for θ_2 and actual market shares S_t trough a contraction mapping we are able to retrieve the estimate of δ_t.

$$\delta_t^{m+1} = \delta_t^m + ln(S_t) - ln(\tilde{\sigma}(\delta_t^m, x_t, p_t, F_{n_s}; \theta_2) \ m = 0, \ldots, M. \tag{5}$$

Where M is the number of iterations required to reach the tolerance level ρ: $\|\delta_t^M - \delta_t^{H-1}\| < \rho$. Recalling that $\delta_{j,t} = \beta^0 + x'_{j,t}\beta^x + \beta^p p_{j,t} + \xi_{j,t}$ we can now formulate the error term using $\tilde{\sigma}^{-1}$:

$$\xi_{j,t}(\theta) = \tilde{\sigma}^{-1}(S_t, x_t, p_t; \theta_2) - \beta^0 - x'_{j,t}\beta^x - \beta^p p_{j,t} \ . \tag{6}$$

The contraction mapping in (5) is called each time the GMM objective function is evaluated, and it represents the inner-loop of BLP algorithm.

To the extent that firms observe the demand shocks ξ_t, and condition on them when they set their prices, the resulting correlation between p_t and ξ_t generates endogeneity and causes bias in the estimates of the parameters. BLP addresses the endogeneity of prices with a vector of instrumental variables $z_{j,t}$. These IVs can be product-specific cost shifters. Usually the non price characteristics in $x_{j,t}$ are also assumed to be mean independent of $\xi_{j,t}$ and hence to be valid instruments. The GMM estimator is then based on the conditional moment condition: $E[\xi_{j,t}|z_{j,t}, x_{j,t}] = 0$. Computationally, the moments are then implemented as $E[\xi_{j,t} \cdot w(z_{j,t}, x_{j,t})] = 0$ for a known function w which generates M moment conditions. A GMM estimator can then be constructed by using the empirical analog of the M conditions:

$$\overline{g}(\xi(\theta)) = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{J}\xi_{j,t}(\theta)\cdot w(z_{j,t},x_{j,t}) \tag{7}$$
$$= \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{J}(\tilde{\sigma}^{-1}(S_t,x_t,p_t;\theta_2) - \beta^0 - x'_{j,t}\beta^x - \beta^p p_{j,t})\cdot w(z_{j,t},x_{j,t})\ .$$

For a weighting matrix W, we define the GMM estimator as the vector θ^* that solves the optimization problem

$$\min_{\theta} g(\xi(\theta))' W g(\xi(\theta))\ . \tag{8}$$

The BLP algorithm consists therefore of an outer-loop to minimize the GMM objective function in (8) and an inner-loop to evaluate this objective function at a given θ by inverting the market share equations.

2 Framework and model specification

2.1 Setting

Hereafter we perform BLP estimation for the market of apps. The main dataset contains data for $T = 36$ different markets built by observing 9 geographies over 3 different years each, yielding a total of 4958 products. We use as characteristics for apps the price, the average score and the availability of in app purchases, therefore

$$\theta^1 = (\beta^0, \beta^p, \beta^{score}, \beta^{iap}) \ .$$

The demographic dataset stores micro data of income and age of 500 consumers for each market considered. In our analysis we will use only income to interact with random coefficients, so that

$$\theta^2 = \begin{pmatrix} \sigma_{RC_1} & \pi^I_{RC_1} \\ \cdots & \cdots \\ \sigma_{RC_N} & \pi^I_{RC_n} \end{pmatrix}$$

where N is the number of random coefficients, π^I is their sensitivity to consumer's income and σ is their relationship with the random component. Among the results, we will analyse the distribution across the markets of the price coefficients $\beta^p_t \ \forall t = 1, \ldots 36$ to check how well BLP models capture price elasticities.

2.2 GMM, IV and heteroskedastic errors

In our BLP model we assume that price p is the endogenous regressor, while the others are exogenous. We generally use only one excluded instrument for price: our matrix of instruments Z would be made of the excluded instrument and the exogenous regressors. Therefore we have a GMM that is exactly identified: the number of parameters to estimate K is equal to the number of moment conditions M we generate trough the instruments. The Generalized Method of Moments will gives us the following moment conditions:

$$E(g(\xi(\theta))) = 0$$
$$\text{where } g(\xi(\theta)) = Z'\xi(\theta)$$

We have thus specified the function $w(z_{j,t}, x_{j,t})$ that was more generally stated in (7). The empirical analog of the moment conditions is

$$\overline{g}(\xi(\theta)) = \frac{1}{TJ} Z'\xi(\theta)$$

A GMM estimator for θ is $\hat{\theta}$ that solves (8), which means solving the K first order conditions (where K is the number of parameters in θ)

$$\frac{\partial g(\xi(\theta))' W g(\xi(\theta))}{\partial \theta} = 0$$

The resulting GMM estimator is

$$\hat{\theta}_{GMM} = (X'ZWZ'X)^{-1}X'ZWZ'S_t \tag{9}$$

In our BLP the GMM function is minimized over the entire θ, made by vertical coefficients θ_1 and consumer specific ones θ_2, therefore X must contain not only the characteristics but also the derivative of shares with respect to the random drawn and to the demographics. Since from (6) there is a one-to one correspondence between market shares and the observable product characteristics ξ_{jt}, the additional columns in X can also be seen as the Jacobian of the unobservables, which are necessary to compute moment conditions.

The efficient GMM estimator is the GMM estimator with an optimal weighting matrix W, one which minimizes the asymptotic variance of the estimator. It can be demonstrated that the optimal W is $V(g)^{-1}$ the inverse of the covariance matrix of the moment conditions g:

$$V(g) = \frac{1}{TJ}E(Z'\Omega Z)$$
$$\text{where } \Omega = E(\xi'\xi)$$

In our BLP framework we assume heteroskedasticity of errors ξ. Denoting by $\hat{\Omega}$ the diagonal matrix of squared errors ξ, a consistent estimator of $V(g)$ is:

$$\hat{V}(g) = \frac{1}{TJ}(Z'\hat{\Omega}Z)$$

In presence of heteroskedasicity, the IV estimator is inefficient but consistent: it does not achieve the lowest variance among all possible estimators, but as the sample size increases, it converges in probability to the true parameter value. In other terms, the IV estimator is a GMM estimator with a sub-optimal weighting matrix $\hat{W} = \hat{V}(g)^{-1}$. With the sub-optimal weighting matrix $\hat{W}$ we obtain an estimated variance-covariance matrix for the IV estimator that is robust to the presence of heteroskedasticty:

$$\hat{V}(\hat{\theta}) = (X'P_Z X)^{-1}(X'Z(Z'Z)^{-1}(Z'\hat{\Omega}Z)(Z'Z)^{-1}Z'X)(X'P_Z X)^{-1} \tag{10}$$

where P_Z is the projection matrix $Z(Z'Z)^{-1}Z'$. In order to estimate standard errors for $\hat{\theta}$ we use (10), which is known as Eicker-Huber-White "sandwich" robust variance-covariance matrix for the IV estimator.

2.3 Optimization and Convergence

The optimization method employed to minimize the GMM function is the Nelder-Mead simplex algorithm, which operates by maintaining a simplex in the parameter space and iteratively adjusting its vertices to minimize the objective function. Nelder-Mead performs a derivative-free optimization: it is able to handle noisy or poorly conditioned objective functions, which are not differentiable. Being a local optimization algorithm, it may converge to a local minimum: therefore multiple runs with different initial guesses may be required to find a global minimum.

In the contraction mapping required to estimate δ_t from a guess of θ^2 and the market shares S_t, the tolerance is set at $\rho = 1e-5$. Moreover a maximum level of iterations for the inner-loop to converge is set at 2500. With respect to the GMM minimization a double convergence threshold is set: the absolute change in θ has to be smaller than $1e-5$ and the change in the objective function $g(\xi(\theta))'Wg(\xi(\theta)$ smaller than $1e-3$. We change the default values of $1e-4$ for both threshold, since the the GMM function is higher in absolute value with respect to the maximum θ coefficient by a factor of $1e3$, therefore asking for a convergence at the same decimal figure would cause a number of iterations in which θ does not minimally change, definitely a waste of resources. For the outer-loop we also set a maximum number of iterations to reach convergence equal to 100.

For each model, we start with an initial guess for the consumer-specific price coefficients $\theta_p^2 = (-0.5, 0.5)$ [1]. We start with a positive prediction for π_p^I since we expect price sensitivity to decrease with respect to income. As the random component σ_p of the price coefficient is by definition quite difficult to predict in sign, we use a value which is exactly the opposite of π_p^I so that, under the assumption that the income distribution is asymptotically normal and that the optimal θ^2 would not be so different from the initial guess, the final distribution of β_h^p is roughly centered on its vertical component β^p.

2.4 Issues in variance-covariance of the coefficients

In all the BLP versions, we encounter issues in the standard errors (SE) of the coefficients: some items of the diagonal of $\hat{V}(\hat{\theta})$ are indeed zero or even negative. We traced the problem back to the inversion of $X'P_ZX$, which happens to be very close to be singular. The BLP code we use already controls for singularity of $X'P_ZX$ and uses Least Squares Solution to approximate its inverse, but is unable to handle with situations where the determinant is very close to zero but not exactly zero. The problem we face is that while $X'P_ZX$ is still technically invertible, the inverse is extremely sensitive to small input changes due to the proximity to singularity.

As suggested by Conlon (2013) the primary source of bias in the BLP estimator $\hat{\theta}$ is the correlation between the moment conditions and their Jacobian, which arises because the unobservable quality ξ_{jt} is correlated with price p_{jt}. Endogeneity makes moments conditions and their empirical analog (7) to be linearly dependent, resulting in a system of moment conditions which is singular or nearly singular. The optimal solution would be to have exogenous cost shifters to use as instruments $z_{j,t}$, but unfortunately in our empirical setting those are not available. Therefore we have to deal with ill-conditioned covariance matrix of the moment vector.

Xiao (2020) demonstrates that the GMM estimator using any of the reflexive generalised inverses, and the Moore-Penrose generalised inverse in particular, as the weighting matrix, will always have the same asymptotic variance as the efficient GMM. Holding with respect to the entire variance matrix of moment conditions, a fortiori we can apply this property when dealing with the simple inverse matrix

[1] For computational reasons, the data on characteristics, demographics and random drawn are scaled by a factor of 100, here as in the following results we present scaled coefficients

$(X'P_ZX)^{-1}$ that we pre- and post- multiply to get the robust estimate (10) . As for the consistency property of our IV estimator, we should not rely on asymptotic properties to sample of small size. Fortunately, in our case the sample is wide enough to ensure a good approximation for the robust IV estimator variance-covariance matrix with pseudo-inverse for $(X'P_ZX)^{-1}$ instead of the inverse.

3 BLP estimation

We estimate four different versions of the BLP model, each one described in the following subsections.

3.1 Group market thickness instrument

The first version of BLP uses as instrument for price the number of apps in the nest, which in our case is a genre classification. It allows only the price coefficient β_h^p to be random, while the other coefficients β^0 and β^x are assumed constant across consumers. Consequently

$$\theta^2 = \theta_p^2 = (\sigma_p \ \pi_p^I) \ .$$

Using one excluded instrument for the only endogenous regressor, we build the matrix of instruments as:

$$Z = (Ngroup \ \ constant \ \ score \ \ iap)$$

where the last three are the exogenous regressors (or included instruments)[2]. As mentioned above, the number of moment conditions M generated by these instruments is equal to the coefficients θ to be estimated: $M = K = 4$. The resulting GMM would then be exactly identified, meaning that the system of equations is exactly determined. In this case the GMM objective function is expected to be precisely zero in correspondence of the optimal estimator θ^*.

3.2 Genre nest

The second version of BLP uses as characteristics also dummy variable for the nest, basically performing a nested logit instead of the simple one of version 1.

After removing the constant, the presence of dummies for genre still induces multicollinearity among regressors. The issue is likely due to the fact that the estimated price relies on exogenous components of variables included in the first stage, among which $Ngroup$ is the only proper instrument, i.e. not included in the second stage. By adding back to this latter stage binary variables for app genres, collinearity trivially obtains. The initial guess for the mean utility δ, performed trough a simple logit, is extremely sensitive to the high condition number of the matrix of characteristics and fixed effects: without handling the problem, the first guess for δ is infinite.

We try different strategies to exit the dummy variables trap: at first we try to group the three fixed effects for the genre of games, which show the highest degree of multicollinearity; then we also drop the dummy variable for the genre "tools" which still shows a high variance inflation factor (VIF). Not yet

[2] A matrix of instruments containing the excluded instrument and exogenous regressors prevents the GMM from being under-identified. A model which lacks sufficient identifying power could lead the estimator not to converge to a unique solution, affects the precision of parameter estimates and makes inference about the true value of the coefficients challenging.

satisfied, we also created a label encoder which assigns different values to a unique column depending on the nest. The VIF analysis for the above mentioned strategies is reported in Table 1, together with the number of extremely high values for δ observed.

Variance Inflation Factor				
		Standard case	**Grouped games**	**Label encoder**
characteristics	price	1.055	1.043	3.461
	average score	1.117	1.101	7.472
	in app purchases	1.185	1.140	1.577
Fixed effects	Education	4.777	1.370	-
	Entertainment	9.731	9.617	-
	Game 1	24.34	-	-
	Game 2	22.20	-	-
	Game 3	25.85	-	-
	Personalization	2.884	2.858	-
	Tools	10.25	10.14	-
	Grouped games	-	69.22	-
	Label encoder	-	-	4.467
Condition number of the matrix		167.9	140.9	15.00
N° of $\delta_{j,t} \geq 1e3$		2084	0	0
of which $\delta_{j,t} \geq 1e6$		1701	0	0

Table 1: VIF for genre nest BLP

While the choice of a label encoder for our genre classification grants no multicollinearity, and consequently a well-posed initial guess for δ, the choice of attributing different values for different genres makes our estimated coefficients θ less easy to interpret. Indeed, we present the results of the use of a label encoder more as a benchmark for absence of multicollinearity than for a real interest in using it. We therefore decide to perform our estimation of the BLP with genre nest using the dataset with grouped games dummies. Our estimated vector of vertical coefficients is:

$$\theta^1 = (\beta^p, \beta^{score}, \beta^{iap}, \beta^{edu}, \beta^{ent}, \beta^{per}, \beta^{too}, \beta^{gam}) \ .$$

Being subject to a certain degree of multicollinearity, even if small, the initial guess for δ with simple logit is slightly upward biased. The inner-loop indeed takes more iterations to converge in this first case, with respect to BLP version 1. From the second call onwards, we notice that the bias disappears and the algorithm is able to converge with a standard number of iterations.

3.3 Hausman instrument

The third version of BLP uses a simple logit as in version 1, but with a different instrument for price: an Hausman style instrument built as the average price of the same app in all other countries. Using such a Hausman-Nevo instrument means relying on the assumption that within the same year, supply shocks are spatially correlated across countries but demand shocks are not. Denoting by C the set of the different countries where app j is sold and by $c \in C$ each country, we can decompose the vertical shock $\xi_{j,c}$ which app j in country c is subject to:

$$\xi_{j,c} = \xi_j^{(1)} + \xi_c^{(2)} + \xi_{j,c}^{(3)} \ .$$

where $\xi_j^{(1)}$ is a common supply shock to app j across all countries where it is sold, $\xi_c^{(1)}$ is a common shock to country c across all the apps sold in that country and $\xi_{j,c}^{(3)}$ is the demand shock specific for that app in that country. Define the instrument for price as

$$Z_{j,c} = \frac{1}{|C| - 1} \sum_{c' \in (C \setminus c)} p_{j,c'}$$

observe that, for the app j observed in country c, its average price across all countries in which it is sold, excluding c, is orthogonal to the app-country specific shock:

$$E[\xi_{j,c}^{(3)} | Z_{j,c}] = 0 \ .$$

Therefore, after controlling for fixed effects, $\xi_j^{(1)}$ and $\xi_c^{(2)}$ our instrument $Z_{j,c}$ is indeed exogenous and does not introduce bias in the estimation.

Now equipped with a new instrument, we also estimate an over-identified GMM with a richer matrix of instruments compared to the previous versions:

$$Z = (Hausman \ \ Ngroup \ \ constant \ \ score \ \ iap) \ .$$

In this case the number of moment conditions is $M = 5$ and it is greater than the parameters to estimate $K = 4$. The system of equations is now over-determined and we expect the GMM objective function to be minimized at a nonzero value, contrarily to exactly identified GMMs. Moreover, the excess of instruments allows to use a Hansen-Sargan test to assess instruments' validity. Results are reported in the Appendix, Table 3. They report a price coefficient estimate in-between those of versions 1,2,4 and that found in version 3 (which relies on the Hausman instrument as the unique exogenous ones), and are generally comparable to the other models in terms of significance and interpretation of coefficients.

3.4 RC on average score

The fourth version of BLP is a simple logit BLP which allows random coefficients also for another characteristics: the average score. Therefore β_h^p and β_h^{score} are consumer specific, while β^0 and β^{iap}

remain vertical. In this case

$$\theta^2 = \begin{pmatrix} \sigma_p & \pi_p^I \\ \sigma_{score} & \pi_{score}^I \end{pmatrix} .$$

In our estimation we start with initial guess

$$\theta^2 = \begin{pmatrix} -0.5 & 0.5 \\ -0.5 & 0.5 \end{pmatrix} .$$

We choose a positive value for π_{score}^I since, assuming that price sensitivity decreases with income, we expect sensitivity to other characteristics to increase. Given that we expect a positive value for the vertical component β^{score} our initial guess aims at enhancing the preference for average score when income increase. As made with price random component, we set σ_{score} at exactly the opposite value of the $\pi_s^I core$ so that, assuming the distribution of income is asymptotically normal and that the optimal θ^2 would not be much different from our initial guess, the final distribution of β_h^{score} is roughly centered on its vertical component β^{score}.

4 Comparative analysis of results

In Table 2 we reported results for the four versions of the BLP models estimated. The output is made of two components: the estimate for vertical parameters θ_1 and that for consumer specific parameters θ_2, with standard errors and the value of the GMM objective function at its minimum.

The reader should note that all characteristics' values were scaled by a factor of 100, thus the interpretation of coefficients must be adjusted by this factor. Consistently with our expectations, across all BLP versions, β^p is always negative and significant. While in versions 1,2,4 it oscillates around -1.5, adding the Hausman instrument reduces its magnitude to -0.38. The average score coefficient is surprisingly not significant in the first two versions, while it becomes positive and significant when paired with the Hausman instrument. Similarly for in-app purchases, the coefficient turns from negative into positive in version 3 (Hausman). Such differences are probably due to the increased price estimate precision with Hausman: estimates in version 3 more closely reflect the sensitivity of demand to each characteristic as collinearity between estimated prices and other characteristics is reduced. This also explains why β^p drops in magnitude that significantly in version 3, as omitted variable bias is reduced.

In version 2 we added group-specific fixed effects. Recall that in our model the constant term (here replaced by fixed effects) represents the value of purchasing the average inside good instead of an outside good. Positive fixed effects suggest that the app categories we included in the analysis are generally preferred to all others (in the outside nest) by consumers.

The consumer-specific component of coefficients is not far from our initial guess. Recall that all estimated BLP versions are exactly identified GMMs: in principle, convergence always occurs and is unique. In practice, however, computational estimation still exposes our estimates to some, even if small, approximation. At the same time, exact identification poses fewer constraints so that the value of the GMM objective function is approximately 0, and convergence is reached considerably quickly, without the need for the algorithm to explore values of θ^2 quite far from our guess. We obtain positive feedback about the accuracy of our guess, as π_p^I always remains significant. If our initial guess were off the tracks, we'd most likely reach convergence still (thanks to exact identification), but π_p^I would not be significant. The same is true for π_{score}^I in version 4. The random component is never significant, which is coherent with our expectations: we do not have consistent paths in the random drawn. Therefore we make the choice of fixing the initial guess for σ_p and σ_{score} to make the coefficients' distributions approximately centered around their vertical components, without affecting results.

Figure 1 reports histograms for the distributions of price coefficients in each BLP version:

$$\beta_h^p = \beta^p + \sigma_p v_h + \pi_p^I D_h$$

In each case, the distribution is right-skewed, and the left tail seems somehow truncated. This phenomenon is mainly due to the distribution of demographics (income in our case). From figure 2 we

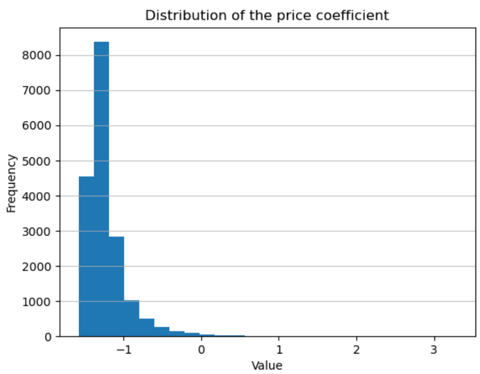
BLP versions comparison					
		N in group instrument	**Genre nest**	**Hausman instrument**	**RC on score**
$\hat{\theta}_1$: characteristics	constant	1.6277	-	-3.1422	3.0071
	(SE)	0.640	-	(0.169)	(0.705)
	price	-122.28	-168.04	-38.148	-151.44
	(SE)	(5.24)	(9.68)	(2.52)	(5.45)
	average score	-8.5335	-7.7440	13.294	-26.510
	(SE)	(8.11)	(9.17)	(3.19)	(4.50)
	in app purchases	-24.126	-32.531	10.519	-31.324
	(SE)	(7.26)	(16.3)	(2.16)	(5.88)
$\hat{\theta}_1$: fixed effects	Education	-	310.74	-	-
	(SE)	-	(15.9)	-	-
	Entertainment	-	505.60	-	-
	(SE)	-	(25.7)	-	-
	Personalization	-	400.42	-	-
	(SE)	-	(30.5)	-	-
	Tools	-	365.60	-	-
	(SE)	-	(9.39)	-	-
	Games	-	343.99	-	-
	(SE)	-	(13.4)	-	-
$\hat{\theta}_2$	price random	-0.52080	-0.51252	-0.53156	-0.48567
	(SE)	(1.94)	(7.07)	(0.309)	(0.646)
	price on income	0.48619	0.50620	0.451683	0.49758
	(SE)	(0.113)	(0.0860)	(0.0275)	(0.134)
	score random	-	-	-	-0.51556
	(SE)	-	-	-	(1.83)
	score on income	-	-	-	0.52471
	(SE)	-	-	-	(0.230)
GMM objective value		1.43e-21	9.62e-21	5.73e-23	1.08e-21
N inelastic demand		185	97	1240	122
% inelastic demand		1.03%	0.539%	6.89%	0.678%

Table 2: Results for BLP estimation

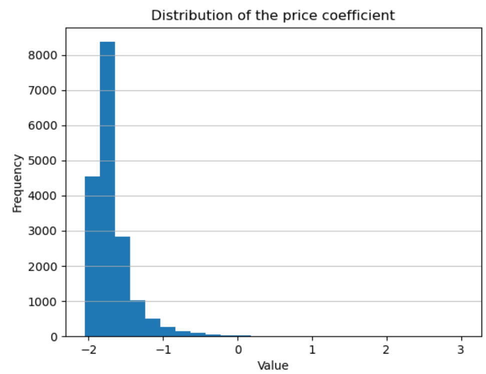
observe that while v_h is normally distributed, the distribution of D_h is inevitably left-bounded, with some observations that are extremely high. Therefore, the choice of a positive initial guess for π_p^I, and a subsequent close-by estimate, yield a long right tail. Above a certain level of income, the negative vertical component β^p, which represents the average sensitivity to price, is offset by the consumer-specific component $\pi_p^I D_h$, yielding a positive price elasticity for certain apps.

Table 2 also reports the number of inelastic demands and their percentage over the entire sample. We observe that in versions 1, 2, 4 the percentage of inelastic demands is around or less than 1%. We could reasonably trace those inelastic demands back to apps whose consumers have outlier income. In version 3, where the Hausman instrument performs better than $Ngroup$, we notice a higher percentage of inelastic demands instead, due to the right shift of the distribution of price elasticites induced by the lower value of vertical component β^p. In this case, more than 6% of inelastic demands for a better instrumented model signals that either there are some apps for which demand is indeed inelastic, or the model specification is missing some characteristic that is relevant for consumers.

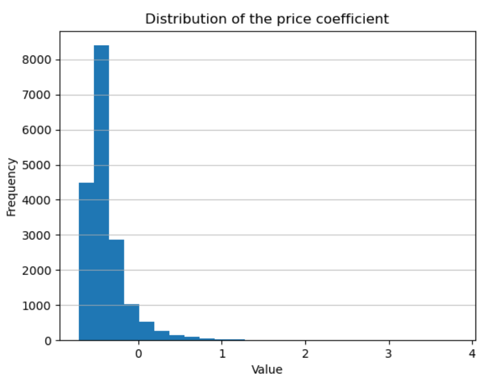
Certainly, with a negative initial guess for π_p^I and a resulting negative estimate $\hat{\pi}_p^I$, the distribution for β_h^p would have been symmetric to the ones obtained in Figure 1 with respect to the vertical component. As a result, no inelastic demands would have arisen. While minimizing the number of inelastic demands is an indicator of a well-specified model, in the latter case it would be impossible to interpret estimated price elasticities. Assuming that price aversion increases with respect to income is very difficult to justify theoretically in our framework, and indeed it would yield to a non-significant $\hat{\pi}_p^I$.



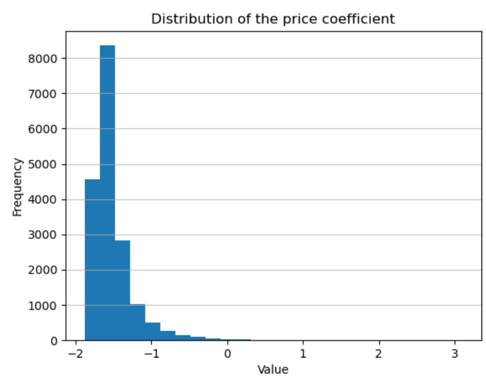
(a) N in group instrument



(b) Genre nest

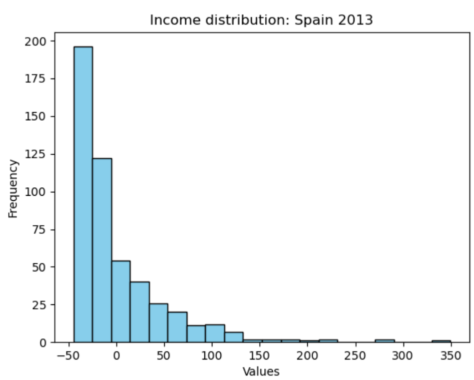


(c) Hausman for price

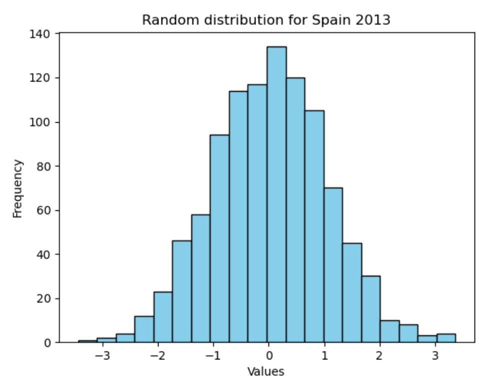


(d) RC on score

Figure 1: Price elasticites



(a) D_h in $T = 1$



(b) v_h in $T = 1$

Figure 2: Data on consumers in Spain 2013

Appendix: Over-identified GMM

Exactly and over-identified BLPs comparison					
		N in group exactly identified	**Hausman exactly identified**	**Hausman & N in group over-identified**
$\hat{\theta}_1$ characteristics	constant	1.6277	-3.1422	-2.6116
	(SE)	0.640	(0.169)	(0.243)
	price	-122.28	-38.148	-60.340
	(SE)	(5.24)	(2.52)	(3.62)
	average score	-8.5335	13.294	12.311
	(SE)	(8.11)	(3.19)	(5.02)
	in app purchases	-24.126	10.519	-8.8917
	(SE)	(7.26)	(2.16)	(5.62)
$\hat{\theta}_2$	price random	-0.52080	-0.53156	-0.87844
	(SE)	(1.94)	(0.309)	(9.89)
	price on income	0.48619	0.451683	0.58438
	(SE)	(0.113)	(0.0275)	(0.0374)
GMM objective value		1.43e-21	5.73e-23	89.1
N inelastic demand		185	1240	934
% inelastic demand		1.03%	6.89%	5.19%
Hansen-Sargan test	J - stat	-	-	0.999
	degrees of freedom	-	-	1
	p - value	-	-	0.317

Table 3: Over-identified BLP

References

Baum, C., Schaffer, M., & Stillman, S., (2002). Instrumental variables and GMM: Estimation and testing. Boston College Economics Working Paper 545

Berry, S., (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2), 242-262

Berry, C., Levinsohn, J., & Pakes, A., (1995). Automobile Prices in Market Equilibrium. *Econometrica*, 63(4), 841-890

Colnon, T., (2013). The Empirical Likelihood MPEC Approach to Demand Estimation *SSRN Electronic Journal*, 2331548

Dubé, J., Fox, J., & Su, C., (2012). Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation. *Econometrica*, 80(5), 2231-2267

Knittel, C., & Metaxoglu, K., (2014). Estimation of random-coefficient demand models: two empiricists' perspective. *The Review of Economics and Statistics*, 96(1), 34-59

Xiao, S., (2020). Efficient GMM estimation with singular system of moment conditions. *Statistical Theory and Related Fields*, 4(2), 172-178