Standard errors in demand estimation with Hausman instruments

Jaewon Lee*

April 11, 2024

Abstract

This paper shows how the common practice of calculating the standard errors of demand estimators can be inaccurate when using Hausman instruments. The typical method ignores a correlation pattern arising from essential endogeneity of Hausman instruments, usually underestimating the true variance of the estimators as a consequence. I explore methods to robustly estimate the variance of the demand estimator for some popular classes of Hausman instruments, including region-based and adjacency-based instruments. Monte Carlo simulations are conducted to evaluate their performances. The results suggest using moderately scoped Hausman instruments and correlation-robust variance estimators, to reduce the true variance of the demand estimator and to accurately estimate the variance.

*I thank my advisors, Philip Haile, Steven Berry, and Yuichi Kitamura, as well as Jonas Lieber and Suk Joon Son for their advice and helpful discussions for this paper.

1 Introduction

Demand models typically require instrumental variables (IVs) for identification and estimation. Several types of instrumental variables have been discussed in the literature, particularly regarding their validity as providing exogenous variations. In contrast, little has been studied in terms of how different classes of instruments may affect the statistical properties of estimators in demand models. In this paper, I examine the impact of the construction of so-called Hausman instruments on the variance of BLP-style estimators. I also identify inaccuracies in conventional methods of variance estimation and propose correct approaches accordingly.

Designed to act as a proxy for exogenous cost shifters, a Hausman instrument is formed as a function of the prices of the same product in other markets. Given that it is a function of equilibrium prices, the Hausman instrument is essentially endogenous in the model. However, the demand shocks in a given market do not affect the equilibrium prices in other markets, provided that demand shocks are independent across markets (although this assumption is often subject to debate in empirical work). Because of this, by construction, the Hausman instrument is uncorrelated with the demand shocks within the market, thus behaving as if it were exogenous, and making the identifying moment condition hold.

This successful separation between demand shocks and endogenous prices, however, breaks down when examining the statistical properties of the demand estimator. The variance of the estimator depends on the variability of the moment condition in the sample, with increased sample moment variation leading to greater estimator variance. It is essential, therefore, to correctly gauge the variability of g_t, a symbol we shall use later to denote deviations of the sample moment from zero.

The common method is to calculate the sample variance of g_t assuming that they are mutually uncorrelated across markets. I show, to the contrary, that these terms can be correlated; when examining the correlation between two markets, say t and t', the demand shocks in t now meet the prices in the same market t. This is because these prices in t appear in market t' as a Hausman instrument. Identifying such correlation patterns is a crucial step in restoring valid inference. The pattern is directly related to the structure of the Hausman instrument in use, whose construction is determined by the researcher, and thus differs from case to case.

I demonstrate how to detect the correlation patterns in general cases and, for some common classes of Hausman instruments, how to obtain the correct standard error in practice. Specifically, for region-based instruments—where markets are grouped into disjoint collections termed *regions*—I recommend calculating the standard errors clustered at the region level. For instruments constructed from adjacent markets, I propose to employ variance estimation techniques from time-series or spatial econometrics, such as those that are robust against autocorrelation or cross-sectional dependence.

I also conduct Monte Carlo simulation exercises to evaluate the performance of the proposed methods. The results suggest that these methods yield more accurate standard errors than conventional approaches that do not account for correlation. In the case of region-based instruments, this relative accuracy remains even in scenarios where the clusters are misspecified as smaller than the actual regions within which the cost shifters are constant.

A caveat is that certain conditions need to be met in order to ensure the consistency of those robust variance estimators. In the context of region-based instruments, the number of clusters must be large. When the instrument is a function of prices across a very large number of markets, such as a national-level Hausman instrument, the number of clusters becomes small by construction. This leads the clustered standard error to perform *worse* than the conventional standard error without correlation correction. However, Hausman instruments with such broad scopes are likely to have high variance. It is therefore recommended to construct more localized Hausman instruments, if possible, and employ correlation-robust variance estimators.

The paper is organized as follows. Section 2 introduces the demand model and the BLP-estimator that are of interest in this paper, and then discusses how the conventional calculation may incorrectly evaluate the estimation accuracy when Hausman instruments are used. Section 3 shows how to obtain the correct standard errors in the aforementioned popular choices of Hausman instruments. Section 4 conducts Monte Carlo simulations to examine the performance of the methods I propose.

2 Hausman IVs and incorrect standard errors

2.1 Discrete choice demand and Hausman instruments

Let consumer i's conditional indirect utility from good $j = 1, \ldots, J$ in market $t = 1, \ldots, T$ be[1]

$$u_{ijt} = x'_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \mu_{ijt} + \epsilon_{ijt}.$$

[1]For simplicity, I assume that the number of inside goods J is the same across markets.

The utility depends on the observed characteristics x_{jt}, the price p_{jt}, and the unobserved (to the researcher) product-level heterogeneity ξ_{jt}. The individual–product-level heterogeneity is represented by μ_{ijt} and ϵ_{ijt}, where μ_{ijt} may be parameterized by γ and may depend on product characteristics, demographic variables, or individual-specific tastes. The idiosyncratic preference shock ϵ_{ijt} is usually assumed to follow type I extreme value distribution. The outside option is represented by $j = 0$ and delivers utility $u_{i0t} = \epsilon_{i0t}$.

The estimator by Berry, Levinsohn, and Pakes (1995, hereafter BLP)—based on the observation by Berry (1994) that the discrete choice model can be estimated by a moment condition—utilizes instrumental variables to address endogeneity caused by the unobserved demand shock ξ_{jt}. Specifically, the (market-level) BLP estimator is a generalized method of moments (GMM) estimator based on the moment condition[2]

$$\mathbb{E}[\xi_{jt}|z_{jt}] = 0,$$

where z_{jt} is a vector of instrumental variables, which are exogenous to the demand shock ξ_{jt} while being relevant to the variations in market shares and prices.

One particular class of instrumental variable of interest in this paper is the so-called Hausman instruments (Hausman, 1996; Nevo, 2000).[3] Designed as a proxy for production cost shifters, a Hausman instrument is a function, usually the average, of the prices of the same product in other markets. The idea is that cost shifters of j may be common (or correlated) in several markets in which j is sold. Then the

[2]In a precise sense, the BLP estimator is a method of simulated moments, as the computation of the moment function usually involves numerical integrations with respect to μ_{ijt}. In this paper, however, I ignore the additional estimation error stemming from the numerical integrations.

[3]Berry and Haile (2021) discuss various classes of instrumental variables that have been proposed in the literature.

common cost shifters may affect the prices of j in those markets. Therefore, even when cost shifters are not observed by the researcher, the prices of j in other related markets may pick up the variation in the common cost shifters of j, and hence serve as a relevant instrument.

To pin down the idea, let the production cost be written as

$$c_{jt} = c_{jt}(w_{jt}, \omega_{jt})$$

where c_{jt} is the marginal cost of the product j in market t, w_{jt} is exogenous cost shifters, and ω_{jt} is idiosyncratic cost shocks.

The cost shifter w_{jt}, if observed, would serve as an instrumental variable under the exogeneity assumption $\mathbb{E}[\xi_{jt}|w_{jt}] = 0$, i.e., the demand shock is mean independent of the cost shifter. Suppose otherwise that w_{jt} is not observed by the researcher, yet he finds another market t' such that $w_{jt'}$ is closely correlated (or coincides) with w_{jt}. Then the researcher can use $p_{jt'}$ as a proxy for w_{jt}, since $p_{jt'}$ is determined as a function of $w_{jt'}$ in equilibrium and the latter is in turn correlated with w_{jt}. Of course, for identification purposes, the instrument must satisfy mean independence, $\mathbb{E}[\xi_{jt}|p_{jt'}] = 0$, which in general requires demand shocks ξ_{jt} and $\xi_{jt'}$ to be independent. I formally introduce this as Assumption 1 below.

2.2 The BLP estimator

Now I show that the commonly calculated standard errors of the BLP estimator can be inaccurate when Hausman instruments are employed, even when the identifying condition is satisfied. To do so, I first describe the estimator and how its standard errors are usually calculated, and then identify the issue arising from this practice.

From now on, consider the following unconditional moment condition

$$\mathbb{E}[z_{jt}\xi_{jt}(\theta_0)] = 0,$$

which is obtained by taking z_{jt} as some function of the original set of instrumental variables and renaming it.[4] The structural error ξ_{jt} is written as a function of the parameter $\theta = (\alpha, \beta, \gamma)$ given the data (including market shares s_t, prices p_t, and product characteristics x_t in market t), following Berry (1994), or Berry, Gandhi, and Haile (2013) for more general cases. Identification requires the moment to be zero at (and only at) the true parameter value θ_0.

With the moment function $g_{jt}(\theta) = z_{jt}\xi_{jt}(\theta)$, the BLP estimator $\widehat{\theta}$ minimizes the criterion function:

$$\widehat{\theta} = \arg\min_{\theta} \overline{g}_n(\theta)' A_n \overline{g}_n(\theta),$$

where $n = JT$ denotes the total number of products across markets, $\overline{g}_n(\theta) = n^{-1}\sum_{jt} g_{jt}(\theta) = n^{-1}\sum_{jt} z_{jt}\xi_{jt}(\theta)$, and A_n is a positive definite weighting matrix.[5]

Under standard regularity conditions, the asymptotic distribution of the estimator is $\sqrt{n}(\widehat{\theta} - \theta_0) \xrightarrow{d} N(0, V)$ as $n \to \infty$, with the asymptotic variance V given by

$$V = (G'AG)^{-1} G'A\Omega AG(G'AG)^{-1}$$

[4] A common choice is to take approximate optimal instruments as suggested by Berry, Levinsohn, and Pakes (1995) and Conlon and Gortmaker (2020), or to construct differentiation instruments by Gandhi and Houde (2019).

[5] For more details and recommended practices, see Conlon and Gortmaker (2020).

where

$$G = \mathbb{E}\left[\frac{\partial}{\partial\theta'} g_{jt}(\theta_0)\right] = \mathbb{E}\left[z_{jt}\left(\frac{\partial}{\partial\theta'}\xi_{jt}(\theta_0)\right)\right],$$

$A_n \xrightarrow{p} A$ as $n \to \infty$, and Ω is the asymptotic variance of

$$\frac{1}{\sqrt{n}}\sum_{jt} g_{jt}$$

where $g_{jt} = g_{jt}(\theta_0) = z_{jt}\xi_{jt}$. The correct estimation of Ω (and hence that of V) is the main objective of this paper, in the context of demand estimation with Hausman instruments.

2.3 The overlooked dependence

The common practice of calculating Ω is to assume implicitly that g_{jt} are uncorrelated across t (while often allowing for heteroscedasticity). Under this assumption, we have $\Omega = \mathbb{E}[g_{jt}g'_{jt}]$ by the central limit theorem with zero correlation, leading to variance estimators that are not robust against correlation between g_{jt}s across t. One such estimator (that is still robust against heteroscedasticity of the demand shock ξ_{jt}) is $n^{-1}\sum_{jt}\widehat{\xi}_{jt}^2 z_{jt}z'_{jt}$, where $\widehat{\xi}_{jt} = \xi_{jt}(\widehat{\theta})$. This is the default variance estimator in PyBLP, a widely used python package, by Conlon and Gortmaker (2020).

When some elements of z_{jt} are Hausman instruments, however, the mutual independence assumption of g_{jt} with respect to t is violated in general. To see this, suppose the price in market t' is used as an instrument for market t and *vice versa*;[6]

[6]Although this is a common practice in empirical work involving Hausman instruments, some settings might not follow such a reciprocal relationship, in which case the correlation phenomenon

i.e., $z_{jt} = p_{jt'}$ and $z_{jt'} = p_{jt}$.[7] Then we have

$$\mathbb{E}[g_{jt}g_{jt'}] = \mathbb{E}[(p_{jt'}\xi_{jt})(p_{jt}\xi_{jt'})] = \mathbb{E}[(p_{jt}\xi_{jt})(p_{jt'}\xi_{jt'})],$$

which is nonzero in general, as an example in Section A shows. This happens because the price p_{jt} of product j and its demand shock ξ_{jt} are correlated, which is in fact why we required instrumental variables in the first place.[8] As p_{jt} and ξ_{jt} are typically positively correlated, the correlation tends to be positive, and ignoring this would underestimate the true variance Ω.

Before discussing various aspects of this phenomenon, here I lay down assumptions on exogenous variables and structural shocks, which I maintain for the rest of the paper.

Assumption 1. (a) Structural errors (ξ_t, ω_t) are independent across markets conditional on exogenous variables $(x, w) = (x_{11}, \ldots, x_{JT}, w_{11}, \ldots, w_{JT})$.

(b) $\mathbb{E}[\xi_{jt}|x, w] = 0$ almost surely.

The first item ensures econometric exogeneity of the Hausman instrument, as demonstrated momentarily. Alternatively, we may consider a set of assumptions in terms of the cost c itself without distinguishing between w and ω, particularly given that w and ω are both unobservable when a researcher resorts to Hausman

does not occur. See below for this.

[7] The vector of instruments z generally contains other elements such as included instruments and other types of instruments, but I suppress them for notational convenience. The discussion for the rest of this section should be interpreted as pertaining to a specific element, namely the Hausman instrument, among the whole vector of instrumental variables.

[8] By the same logic, should we choose to use product shares in other markets as an instrumental variable, we would again have nonzero correlation. This is because the equilibrium shares are endogenous in their respective markets.

instruments; i.e., we may assume $\xi_t \perp\!\!\!\perp \xi_{t'}|x,c$ and $\mathbb{E}[\xi_{jt}|x,c] = 0$. However, keeping ω separate from w is more general in that it allows some part of the cost to be correlated with the demand shock.

Instrument exogeneity/relevance does not prevent nonzero correlation. The nonzero correlation occurs even though econometric exogeneity of the Hausman instrument is maintained. Under Assumption 1, the moment condition holds as

$$\mathbb{E}g_{jt} = \mathbb{E}[p_{jt'}\xi_{jt}] = \mathbb{E}[p_{jt'}\mathbb{E}[\xi_{jt}|p_{t'},x,w]] = \mathbb{E}[p_{jt'}\underbrace{\mathbb{E}[\xi_{jt}|x,w]}_{=0}] = 0,$$

where the third equality holds because $p_{t'}$, once conditioned on exogenous variables, is a function of $(\xi_{t'},\omega_{t'})$, which are conditionally independent of ξ_{jt}. The display guarantees the instrument validity of $p_{jt'}$ for ξ_{jt} as needed. However, under the same set of assumptions, we have

$$\mathbb{E}[g_{jt}g_{jt'}] = \mathbb{E}[\underbrace{\mathbb{E}[p_{jt}\xi_{jt}|x,w]}_{\neq 0}\underbrace{\mathbb{E}[p_{jt'}\xi_{jt'}|x,w]}_{\neq 0}] \neq 0,$$

in general, resulting in a nonzero correlation. Here, the first equality holds again because the structural errors are conditionally independent across markets. Applying the usual modeling implication that $\mathbb{E}[p_{jt}\xi_{jt}|x,w] > 0$ almost surely, we can further say that the sequence of g_{jt} exhibits positive correlation across t in this case.

The nonzero correlation is not an implication of the common/correlated cost shifter assumption either, which is motivated by the relevance requirement of Hausman instruments. This can be seen from the fact that whether w_t is correlated with

$w_{t'}$ or not plays no role in the derivation above, since $\mathbb{E}[p_{jt}\xi_{jt}|x,w]$ is already conditioned on w.

Construction of Hausman instrument determines the correlation pattern. In the above exposition, I constructed the Hausman instrument in a reciprocal way; $p_{t'}$ serves as an instrument for market t and p_t as an instrument for market t'. While such a reciprocal construction of Hausman instruments is common in empirical studies, we can consider otherwise. For example, suppose, for each t, a researcher lets $p_{j,t-1}$ instrument for market t, and *not vice versa*. Then we instead have

$$\mathbb{E}[g_{jt}g_{j,t-1}] = \mathbb{E}[\mathbb{E}[p_{j,t-1}\xi_{j,t-1}|x,w]\mathbb{E}[p_{j,t-2}|x,w]\underbrace{\mathbb{E}[\xi_{jt}|x,w]}_{=0}] = 0,$$

and similarly for other correlations such as $\mathbb{E}[g_{jt}g_{j,t-2}] = 0$.

In the same way, the correlation problem does not arise when prices in t' are used as an instrument for market t while t' itself is not included in the sample. For example, suppose a researcher has data (including prices, market shares, and product characteristics) for the capital of each state, while for other cities he only observes prices. For each capital city he can build an instrumental variable using the prices in other cities in the same state, and then calculate the sample moment function $\overline{g}_n(\theta)$ using only the capital cities. In this scenario, g_{jt}s are uncorrelated under the assumption that demand shocks are independent across states conditional on exogenous variables, thus eliminating the correlation problem.

Note that such patterns of correlation are determined by the researcher's own decision on which markets enter the Hausman instrument for another market, *not* by the true correlation among the cost shifters across markets. Therefore one needs to examine these patterns based on his own choice, on a case-by-case basis, to accurately estimate Ω. In the next section, I explore a few common and popular cases, namely region-based and adjacency-based Hausman instruments. In more general cases, researchers can follow the preceding demonstration to identify the correlation pattern and then adapt robust variance estimators accordingly.

Other instruments do not introduce the hidden correlation. If Hausman instruments that utilize prices in nearby markets introduce correlation, then how about so-called BLP instruments that use the exogenous characteristics of closely related products, or Waldfogel instruments that rely on the exogenous characteristics (e.g., consumer demographics) of nearby markets? What if a researcher directly observes the cost shifters and employ them as instruments? It turns out that these instruments do not face the hidden correlation phenomenon exhibited above, because they are essentially exogenous. Still, one might need to account for correlations between demand shocks, which is usually modeled in an explicit way and therefore easier to detect.

The BLP instruments use exogenous variables x_t (often within the same market) as their ingredients. Reflecting this, now suppose $z_{jt} = x_{j't}$ and $z_{j't} = x_{jt}$. Then we have

$$\mathbb{E}[g_{jt}g'_{j't}] = \mathbb{E}[(x_{j't}\xi_{jt})(x_{jt}\xi_{j't})'] = \mathbb{E}\left[x_{jt}x'_{j't}\mathbb{E}[\xi_{jt}\xi_{j't}|x]\right],$$
$$\mathbb{E}[g_{jt}g'_{jt'}] = \mathbb{E}[(x_{j't}\xi_{jt})(x_{j't'}\xi_{jt'})'] = \mathbb{E}\left[x_{j't}x'_{j't'}\mathbb{E}[\xi_{jt}\xi_{jt'}|x]\right].$$

Whether these expressions evaluate to zero depends on the assumption on how demand shocks are (conditionally) dependent on each other. If the researcher is willing to assume that demand shocks are independent, then the correlations are zero. If, otherwise, he believes that demand shocks within a market are correlated, or that the demand shocks of a product persist across markets, then the corresponding correlations between g_{jt}s are nonzero in general. Such correlation patterns among g_{jt} can be directly deduced, like this, from the assumed correlation among ξ_{jt}s, making it better understood in the literature; in empirical studies, researchers often report standard errors clustered at market-level or product-level, accounting for their assumed correlation between demand shocks. This correlation between g_{jt}s, however, is distinct from the hidden pattern I have revealed as above, in that it is a direct implication of the true demand shocks, rather than an artifact of how an instrument is constructed by the researcher.

We can apply the same logic to Waldfogel instruments. Letting x_t incorporate relevant exogenous variables, such as consumer demographics, we can write the Waldfogel instrument for product j in market t as $z_{jt} = x_{t'}$. Then

$$\mathbb{E}[g_{jt}g'_{j't}] = \mathbb{E}[(x_{t'}\xi_{jt})(x_{t'}\xi_{j't})'] = \mathbb{E}\left[x_{t'}x'_{t'}\mathbb{E}[\xi_{jt}\xi_{j't}|x]\right],$$
$$\mathbb{E}[g_{jt}g'_{jt'}] = \mathbb{E}[(x_{t'}\xi_{jt})(x_{t}\xi_{jt'})'] = \mathbb{E}\left[x_{t'}x'_{t}\mathbb{E}[\xi_{jt}\xi_{jt'}|x]\right],$$

and whether they are zero or not depends on the assumptions on the correlations between demand shocks. Likewise, if the researcher has an access to the exogenous cost shifters, then a similar algebra leads us to

$$\mathbb{E}[g_{jt}g'_{j't}] = \mathbb{E}[(w_{jt}\xi_{jt})(w_{j't}\xi_{j't})'] = \mathbb{E}\left[w_{jt}w'_{j't}\mathbb{E}[\xi_{jt}\xi_{j't}|w]\right],$$
$$\mathbb{E}[g_{jt}g'_{jt'}] = \mathbb{E}[(w_{jt}\xi_{jt})(w_{jt'}\xi_{jt'})'] = \mathbb{E}\left[w_{jt}w'_{jt'}\mathbb{E}[\xi_{jt}\xi_{jt'}|w]\right].$$

The key distinction between Hausman instruments and other types of instrumental variables—such as BLP instruments, Waldfogel instruments, and cost shifters—lies in their endogeneity within the demand model. Hausman instruments, derived from equilibrium prices, are intrinsically endogenous (although they are treated as exogenous for econometric purposes). In contrast, other instruments are exogenous to the model. Consequently, analyzing the correlation between instrumental variables and demand shocks was required for Hausman instruments, but not for the other cases.

3 Correct standard errors

Fortunately, correcting the standard errors becomes straightforward once we identify the dependence structure as in the previous section, since relevant econometric tools are readily available. In this section, I propose the correct calculation of standard errors under two popular classes of Hausman instruments: region-based instruments and adjacency-based instruments. A takeaway is to consider robust variance estimators that account for the dependence pattern, which relies on the particular choice of Hausman instrument.

3.1 Clustered markets

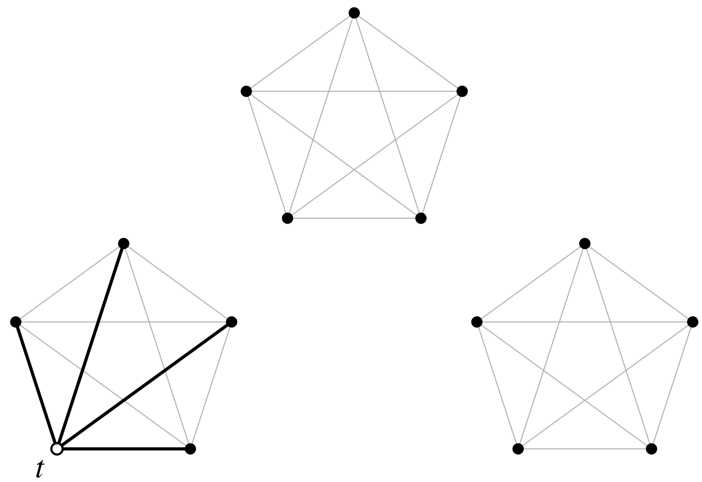
In some studies the Hausman instrument is constructed based on regions, where each region consists of multiple markets. For each market, the researcher constructs a Hausman instrument as a function of the prices in the other markets belonging to the same region. One extreme case is to consider the entire set of markets as a single region: e.g., the Hausman instrument at the national level. For example, Nevo (2000), who defines a city–quarter pair as a market, mentions both approaches—constructing a Hausman instrument within region–quarter clusters or using the prices of the same product across all cities and quarters (except for the same city–quarter pair)—although only the former is used.

Figure 1a illustrates one example, where there are three regions, each composed of five cities. A faint line connecting two markets indicates that one market is used for constructing the Hausman instrument for the other market, and vice versa. For market t, indicated by a hollow dot, the four markets connected by solid lines in the same region are incorporated into the instrument for this market.

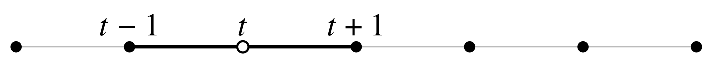
In such cases, the correlation of g_{jt} shows a clustered pattern. To see this, let t denote markets and r denote regions which are disjoint sets of markets. For market t, let $r(t)$ denote the region that t belongs to. The Hausman instrument z_{jt} for product j in market t is then constructed as a function of p_{js} with $s \in r(t) \setminus \{t\}$. Then, for t and t' with $r(t) = r(t')$, we have

$$\mathbb{E}[g_{jt}g_{jt'}] = \mathbb{E}[(z_{jt}\xi_{jt})(z_{jt'}\xi_{jt'})] = \mathbb{E}[(z_{jt'}\xi_{jt})(z_{jt}\xi_{jt'})] \neq 0$$

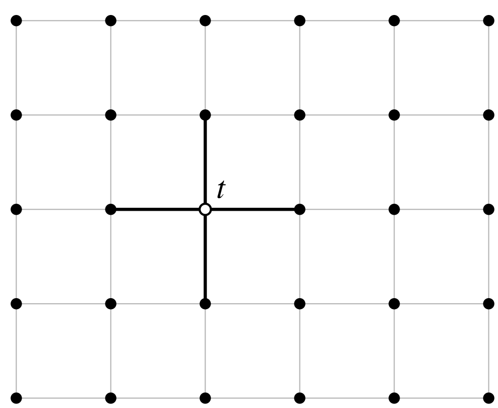
in general, as demonstrated in the previous section, since $z_{jt'}$ depends on p_{jt} which



(a) Clustered markets



(b) Markets on a line



(c) Markets on a lattice

Figure 1: Graphical illustration of markets

Notes: A dot represents a market. A line connecting two markets indicates that one market is included in the Hausman instrument for the other market and vice versa. The reference market t is represented by a hollow dot. The Hausman instrument for t is constructed using the markets that are connected to t by solid lines.

in turn depends on ξ_{jt} in equilibrium, and similarly for z_{jt} and $p_{jt'}$. On the other hand, if t and t' are such that $r(t) \neq r(t')$, then $g_{jt} = z_{jt}\xi_{jt}$ is a function of $(\xi_s, \omega_s, x_s, w_s)_{s\in r(t)}$ whereas $g_{jt'}$ is a function of $(\xi_s, \omega_s, x_s, w_s)_{s\in r(t')}$, so that g_{jt} and $g_{jt'}$ become independent conditional on exogenous variables:

$$\mathbb{E}[g_{jt}g_{jt'}] = \mathbb{E}[\mathbb{E}[g_{jt}g_{jt'}|x, w]] = \mathbb{E}[\mathbb{E}[g_{jt}|x, w]\mathbb{E}[g_{jt'}|x, w]] = 0.$$

Based on this, we can calculate the clustered standard error as follows, by treating each region–product pair as a cluster:

$$\widehat{\Omega} = \frac{1}{n}\sum_{j,r} g_{jr}g_{jr}' \tag{1}$$

where $g_{jr} = \sum_{t\in r}\widehat{\xi}_{jt}z_{jt}$. This clustering scheme is valid under the assumption that demand shocks are (conditionally) independent across products within a market. As this might be restrictive, we may instead cluster at the region level, considering products within a region as belonging to the same cluster, to allow for correlation within each market:[9]

$$\widehat{\Omega} = \frac{1}{n}\sum_{r} g_{r}g_{r}' \tag{2}$$

where $g_r = \sum_j \sum_{t\in r}\widehat{\xi}_{jt}z_{jt}$. In either case, it is convenient to implement in practice since clustered standard errors are built in as an option in many statistical packages including PyBLP.

[9]In fact, this unnecessarily allows for other correlations, say, between g_{jt} and $g_{j't'}$ with $r(t) = r(t')$.

An important consideration, however, is the number of clusters. For the sake of clarity in the following explanation, assume that each region contains an equal number of markets, denoted as $|r|$, and that each region–product pair is treated as a cluster for constructing the Hausman instrument. Consequently, the number of observations in each cluster, which we refer to as the *scope* of the instrument, equals $|r|$. Then given the total number of products across all markets $n(\equiv J \times T)$, the number of clusters equals $n/|r|$.

Given the set of markets, a researcher decides on the scope of the Hausman instrument, and consequently, the number of clusters. Econometric theory prescribes that the number of clusters should be large to ensure consistency of the clustered variance estimator $\widehat{\Omega}$. This requirement presents a practical trade-off; on one hand, the estimation of the true variance becomes more accurate when the number of clusters is large. Also, an excessively broad scope (and thus having too few clusters) may weaken the instrument, by including irrelevant markets and by generating little variation in the instrument. On the other hand, relying on too few markets as proxies for production cost may also weaken the instrument. As demonstrated by Monte Carlo simulations in the next section, I recommend finding moderately sized regions in which cost shifters are believed to be common—to reduce the true variance—and applying clustered standard error—to accurately estimate the true variance.

3.2 Adjacent markets

Another possibility is to build Hausman instruments based on adjacent or nearby markets, rather than partitioned regions. For example, Guevara and Ben-Akiva (2006) use "observed average prices of the same product in adjacent zones," although their estimation method is different from BLP. In such cases, g_{jt}s exhibit serial/spatial correlation, calling for variance estimation techniques from time series, spatial, or network econometrics.

Suppose, for simplicity, markets $t = 1, \ldots, T$ are placed on a line as in Figure 1b. For each t, the instrumental variable is based on the prices of the same product in the two adjacent markets (or one if the market of interest is at either end of the line). Then we can verify that $\mathbb{E}[g_{jt} g'_{j,t-1}] \neq 0$ in general, whereas $\mathbb{E}[g_{jt} g'_{j,t-\ell}] = 0$ for $\ell > 1$, exhibiting a serial correlation similar to an $MA(1)$ process in time series. Analogous patterns arise when more markets are captured by the Hausman instrument. Note that the order of serial correlation (one in this specific case) depends on the choice of the Hausman instrument, not on the true correlation structure of the cost shifters.

Autocorrelation-robust variance estimators can be employed for such cases. A popular choice is by Newey and West (1987). Given the maximum lag L chosen by the researcher, the estimator for Ω is

$$\widehat{\Omega} = \frac{1}{n} \sum_{j} \left[\sum_{t=1}^{T} \widehat{\xi}_{jt}^{2} z_{jt} z'_{jt} + \sum_{\ell=1}^{L} \sum_{t=\ell+1}^{T} \kappa_{\ell} \widehat{\xi}_{jt} \widehat{\xi}_{j,t-\ell} \left(z_{jt} z'_{j,t-\ell} + z_{j,t-\ell} z'_{jt} \right) \right]$$

where κ_{ℓ} is the Bartlett kernel:

$$\kappa_{\ell} = \begin{cases} 1 - \dfrac{\ell}{1+L} & \ell \leq L \\ 0 & \ell > L. \end{cases}$$

The first term in $\widehat{\Omega}$ forms the usual heteroscedasticity-robust variance estimator. The second term adds autocovariances between g_{jt}s across markets. Although the order of serial correlation is known by the construction of the Hausman instrument, the literature nevertheless recommends setting $L \to \infty$ as $T \to \infty$.[10] A rule-of-thumb choice is $L = \lfloor 0.75T^{1/3} \rfloor$.

In more realistic scenarios where cities are situated in a two-dimensional space or a network, methods from spatial or network econometrics can be employed depending on the assumption about how markets are related to each other and on the construction of Hausman instruments. Conley (1999) introduces a spatial model of dependence and provides a consistent variance estimator. For instance, if markets $(s,t) \in \{1,\dots,S\} \times \{1,\dots,T\}$ are in an integer lattice as in Figure 1c, the estimator therein can be written as

$$\begin{aligned}\widehat{\Omega} = \frac{1}{n}\sum_j \left[\sum_{k=0}^{L_S}\sum_{\ell=0}^{L_T}\sum_{s,t} \kappa_{k\ell}\widehat{\xi}_{j,(s,t)}\widehat{\xi}_{j,(s-k,t-\ell)}\big(z_{j,(s,t)}z'_{j,(s-k,t-\ell)} + z_{j,(s-k,t-\ell)}z'_{j,(s,t)}\big)\right] \\ - \frac{1}{n}\sum_j\sum_{s,t}\widehat{\xi}^2_{j,(s,t)}z_{j,(s,t)}z'_{j,(s,t)},\end{aligned}$$

with suitable choices of maximum lags L_S and L_T, where

$$\kappa_{k\ell} = \left(1 - \frac{|k|}{1+L_S}\right)\left(1 - \frac{|\ell|}{1+L_T}\right) \cdot \mathbb{1}\{|k| \le L_S, |\ell| \le L_T\}.$$

As an extension of the Newey–West estimator to multidimensional Euclidean spaces, this variance estimator also accounts for covariances between g_{jt}s across markets.

We may consider various other estimation methods depending on the inter-market structure in each particular empirical context. However, its discussion is beyond the scope of this paper. The key point is that whenever the price in t' is incorporated to form a Hausman instrument for t (and vice versa), the dependence

[10] See Section 6.2.2 of Ogaki (1993). The main rationale is to ensure positive definiteness of $\widehat{\Omega}$.

between t and t' must be taken into account when calculating the standard error.

4 Monte Carlo simulation

In this section, I conduct Monte Carlo simulations to evaluate the performance of standard errors that account for the correlation pattern induced by two popular classes of Hausman instruments: region-based and adjacency-based instruments.

From now on, by *non-robust standard errors* I mean standard errors that are only robust against heteroscedasticity, i.e., the default of PyBLP. By *robust standard errors* I mean standard errors that correctly account for the correlation between g_{jt}s. In the clustered markets case, the robust standard error is the clustered standard error. In adjacent markets case, it is the standard error methods by Newey and West (1987) and Conley (1999).

4.1 The data generating process

The data is simulated according to the nested logit model where the indirect utility from product j belonging to group g is given by

$$u_{ijt} = \beta_0 + \beta_1 x_{jt} - \alpha p_{jt} + \xi_{jt} + \zeta_{ig} + (1-\gamma)\epsilon_{ijt}$$

where γ is the nesting parameter. I set the true parameter values as $\alpha = 1$, $\beta = (\beta_0, \beta_1) = (1, 1)$, $\gamma = 0.5$, and hence $\theta = (\alpha, \beta, \gamma) = (1, 1, 1, 0.5)$.

The product characteristic x_{jt} is drawn from the standard normal distribution. The preference shock ϵ_{ijt} is independently drawn from the type I extreme value distribution, and the individual-group specific shock ζ_{ig} is drawn from the unique distribution that makes $\zeta_{ig} + (1 - \gamma)\epsilon_{ijt}$ an extreme value random variable.[11]

The marginal cost is specified as

$$c_{jt} = 1 + w_{jt} + \omega_{jt}.$$

How I draw w_{jt} will vary in each setting as described below. Demand and cost shocks ξ_{jt} and ω_{jt} are drawn from a bivariate normal distribution, each with a mean of zero and a variance of two, their correlation coefficient being 0.9. They are drawn independently across t and j and independently from other variables.

There are $T = 600$ markets, each having two groups for inside goods. Each group consists of two inside goods ($J = 4$ in total). Each product is produced by a distinct single-product firm. The equilibrium prices and market shares are determined in the Bertrand–Nash price-setting game.

I estimate the demand parameters via two stage least squares, using Hausman instruments that are the average price of the same product in other markets within the scope of the instruments. The scope—i.e., the markets captured by a Hausman instrument (or the number of such markets depending on the context)—varies in each setup. In addition to the Hausman instrument, I use two BLP instruments: the sum of the characteristic of the rival products in the same market and the sum of the characteristic of the rival product within the same group.

[11] See Berry (1994) and Cardell (1997) for more details on the nested logit model.

4.2 Clustered markets

I first set the scope of the Hausman instrument to correctly follow the *actual* region (in which the markets share the same cost shifter as per the data generating process), and examine the performance of the standard errors. The common cost shifter is drawn from the standard normal distribution for each region.

The results are tabulated in Table 1. The standard errors for $\widehat{\alpha}$ are plotted in Figure 2.[12] The true standard error of the price coefficient estimator (simulated by Std. dev. $\widehat{\alpha}$) initially declines with increasing region size, likely because the Hausman instrument gathers more information on the cost. However, as the regions become very large, the true standard error of $\widehat{\alpha}$ starts to grow. This phenomenon may be due to excessive noise entering the instrument, or due to little variation in the instrumental variable across markets.

Overall, the non-robust standard error is inaccurate; the actual coverage of the nominal 95% confidence interval based on the non-robust standard error is lower than 95%, meaning that the non-robust standard error overestimates the accuracy of the demand estimate. On the other hand, the robust standard error performs better than the non-robust one until the size of regions reaches 50. After that, the performance of the robust standard error deteriorates likely due to the insufficient number of clusters. For small number of clusters, the non-robust standard error seems to be a better choice.

The results for other parameters are mostly omitted from the table. I find that $\widehat{\beta}_0$ exhibits a pattern similar to $\widehat{\alpha}$. Surprisingly, the behaviors of $\widehat{\beta}_1$ and $\widehat{\gamma}$ remain stable

[12]The region size 600 is omitted from the figure due to its disproportionately large magnitude on the y-axis; the standard deviation of $\widehat{\alpha}$ for this case is 0.258. Similar omissions occur in Figure 3. The x-axis is not to scale in this and subsequent figures.

Table 1: Standard errors in clustered markets (correctly specified regions)

	α								
Size of a region	2	3	6	12	24	60	120	300	600
# clusters	1200	800	400	200	100	40	20	8	4
Average $\widehat{\theta}$	1.016	1.003	1.005	1.004	1.002	1.003	1.005	1.005	0.998
Std. dev. $\widehat{\theta}$	0.142	0.103	0.083	0.072	0.065	0.066	0.067	0.080	0.258
Robust s.e.	0.138	0.103	0.080	0.070	0.064	0.062	0.062	0.061	0.118
95%-coverage	0.947	0.947	0.945	0.946	0.948	0.933	0.918	0.851	0.751
Non-robust s.e.	0.110	0.087	0.072	0.066	0.063	0.064	0.066	0.074	0.144
95%-coverage	0.900	0.907	0.912	0.933	0.944	0.943	0.949	0.944	0.949
	β_0			β_1			γ		
Size of a region	2	24	600	2	24	600	2	24	600
Average $\widehat{\theta}$	1.017	1.002	0.998	1.001	0.999	0.999	0.500	0.500	0.501
Std. dev. $\widehat{\theta}$	0.149	0.077	0.304	0.061	0.059	0.065	0.043	0.042	0.045
Robust s.e.	0.149	0.076	0.142	0.061	0.059	0.053	0.044	0.042	0.039
95%-coverage	0.959	0.943	0.702	0.955	0.950	0.822	0.956	0.950	0.844
Non-robust s.e.	0.121	0.075	0.170	0.061	0.059	0.064	0.043	0.043	0.045
95%-coverage	0.907	0.945	0.944	0.955	0.955	0.953	0.954	0.955	0.952

Notes: Each experiment is based on 2,000 simulation draws. Average $\widehat{\theta}$ is the sample mean of the simulated estimates, and Std. dev. $\widehat{\theta}$ is the standard deviation of the estimates. Robust s.e. is the mean of the estimated standard errors that are robust against correlation between g_{jt}s. Non-robust s.e. is the mean of the estimated standard errors that are only robust against heteroscedasticity. 95%-coverage calculates the actual coverage of a nominal 95% confidence interval formed by using the respective standard error.

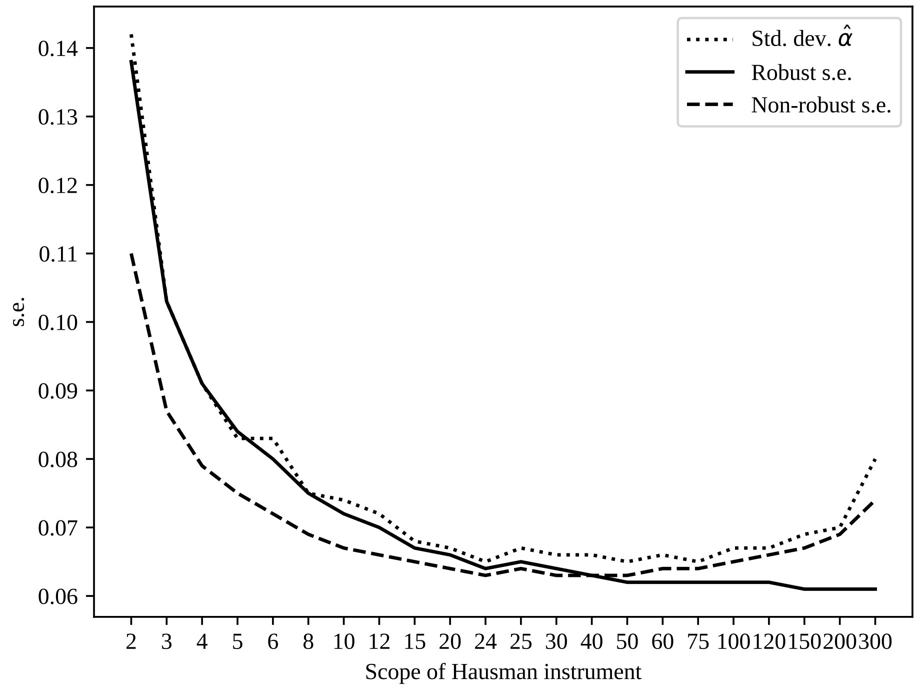


Figure 2: Standard errors for $\widehat{\alpha}$ in clustered markets (correct regions)

Notes: The x-axis is the scope of the Hausman instrument, i.e., the size of each cluster used in the clustered standard error. The curves are plotted under the assumption that the actual regions (in which the cost shifter is constant across markets) coincide with the clusters.

across various region sizes, with their true standard errors consistently low, ranging between 0.04 and 0.06. Both robust and non-robust standard errors accurately estimate the true standard errors—except in cases of very large regions, possibly due to contamination caused by poor estimation of α and β_0. It appears that BLP instruments are sufficient to identify these parameters in this special case of the nested logit model. However, this is not the case in general.[13]

Next, I consider the performances of standard errors under various scopes of

[13]For identification in general nonparametric settings, see Berry and Haile (2014).

the Hausman instrument, while keeping the *actual* region size at 12. This exercise explores the impact of misspecifying the scope of the common cost shifter. Results are shown in Table 2 and Figure 3.[14] Note that, in each case, there are 50 actual regions, each with 12 markets (i.e., the size of actual regions is 12) sharing a common cost shifter. "Scope of IV = 2" refers to the case in which the researcher divides each actual region of size 12 further into 6 clusters of size 2 (hence "scope of IV = 2") to construct the Hausman instrument. On the other hand, "Scope of IV = 24" involves merging two regions into one cluster consisting of 24 markets to construct the instrument.

The true standard error (denoted as Std. dev. $\widehat{\alpha}$) of the estimator is the smallest when the researcher correctly specifies the scope within which the cost shifter is shared. Including too few markets provides not enough proxies for the cost, whereas including too many markets makes the instrument weak. The robust standard error performs better than the non-robust one in most cases. The robust standard error approximates the true standard error particularly well when the researcher sets the scope of the Hausman instrument smaller than the actual region. It does not perform better than the non-robust standard error with large scopes, likely due to the small number of clusters. Both methods show very poor performance when the scope is very large. This is probably partly due to the poor performance of the parameter estimator $\widehat{\theta}$ under weak instrumental variable.

Similar results can be found in Appendix B, where the size of the actual regions is varied over 12, 30, 60, 120, 300, and 600. The robust standard error consistently

[14]Only the results for the price coefficient α are listed, as other parameters exhibit essentially the same pattern as before; the estimate for β_0 behaves similarly to that of α, and the estimates for β_1 and γ are not affected by the choice of the Hausman instrument.

Table 2: Standard errors for $\widehat{\alpha}$ in clustered markets (misspecified regions)

Scope of IV	2	3	4	6	12	24	60	120	300	600
# clusters	1200	800	600	400	200	100	40	20	8	4
Average $\widehat{\alpha}$	1.009	1.005	1.006	1.004	1.004	1.010	1.027	1.047	0.957	0.718
Std. dev. $\widehat{\alpha}$	0.138	0.105	0.093	0.081	0.072	0.096	0.173	0.390	1.060	1.158
Robust s.e.	0.137	0.104	0.092	0.080	0.070	0.093	0.148	0.273	0.934	1.162
95%-coverage	0.938	0.950	0.953	0.951	0.946	0.947	0.933	0.913	0.827	0.628
Non-robust s.e.	0.110	0.088	0.080	0.072	0.066	0.088	0.145	0.272	0.972	1.396
95%-coverage	0.894	0.907	0.916	0.920	0.933	0.934	0.936	0.926	0.877	0.734

Notes: Each experiment is based on 2,000 simulation draws. Average $\widehat{\alpha}$ is the sample mean of the simulated estimates, and Std. dev. $\widehat{\alpha}$ is the standard deviation of the estimates. Robust s.e. is the mean of the estimated standard errors that are robust against correlation between g_{jt}s. Non-robust s.e. is the mean of the estimated standard errors that are only robust against heteroscedasticity. 95%-coverage calculates the actual coverage of a nominal 95% confidence interval formed by using the respective standard error. In all columns, the size of the actual region is 12.

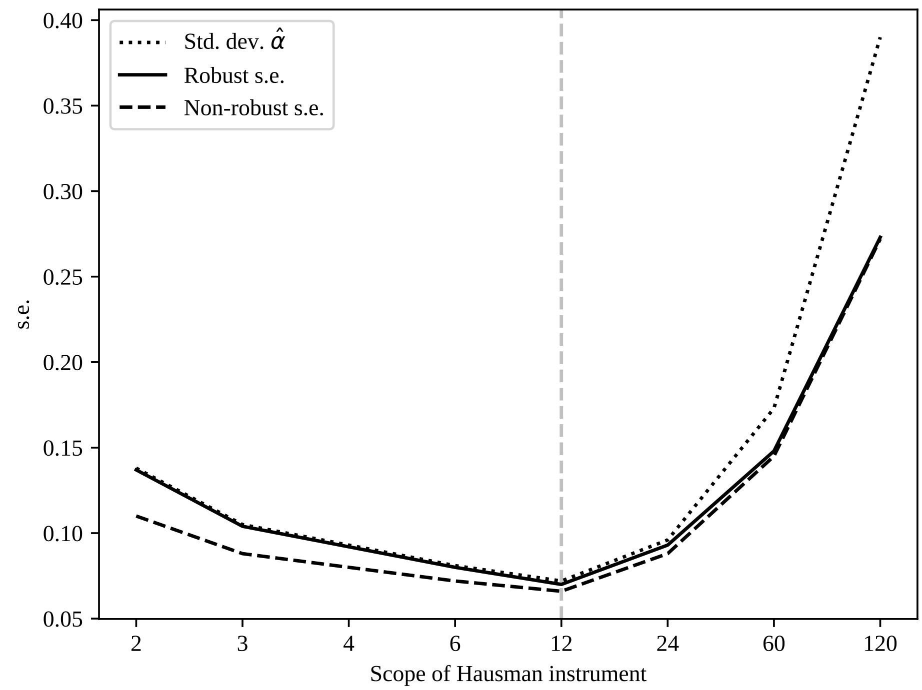


Figure 3: Standard errors for $\widehat{\alpha}$ in clustered markets (misspecified regions)

Notes: The x-axis is the scope of the Hausman instrument, i.e., the size of each cluster used in the clustered standard error. A vertical dashed line is positioned to denote the size of the actual region, which is 12.

outperforms when the scope of the instrumental variable is small. However, with fewer clusters, it performs worse than the non-robust standard error.

To summarize, the true variance is relatively small when the size of actual regions is moderate. The robust standard error is accurate when the number of clusters is large. It is therefore recommended to find moderately sized regions within which some cost shifters are presumed to be common, and apply the clustered standard error. In situations where one must rely on the assumption of a national-level cost shifter, then the non-robust standard error might be a better choice.

4.3 Adjacent markets

Two cases are considered in this subsection, in which markets are on a line and on an integer lattice, respectively. In both cases, the Hausman instrument is the average of the prices of the same product in the immediately adjacent markets, as in Figures 1b and 1c. Therefore, the Hausman instrument captures two markets in the former case, and four markets in the latter case.

In the line world, the cost shifter w_t is generated as the moving average of five normal variates so that the variance of w_t is one, i.e., $w_t = (\eta_{t-2}+\eta_{t-1}+\cdots+\eta_{t+2})/\sqrt{5}$ with each η_t drawn independently from the standard normal distribution. In the lattice world where each market is represented by a pair (s, t), the cost shifter w_{st} is generated as the moving average of 25 normal variates such that the variance of w_{st} is one, i.e., $w_{st} = \sum_{\ell=-2}^{2} \sum_{k=-2}^{2} \eta_{s+\ell,t+k}/5$, where each η_{st} is an independent standard normal variate.

I use variance estimators by Newey and West (1987) and Conley (1999), respectively. Unlike their original forms, however, I let $L = 1$ and use $\kappa_1 = 1$ instead for the Newey–West estimator, and analogously for the Conley estimator. Using weights $\kappa_\ell < 1$ and letting $L \to \infty$ was originally motivated by the anomaly that the calculated variance matrix may fail to be positive definite, but I do not observe this phenomenon in the current simulation exercise.

Table 3 tabulates the results, which are similar to what we have obtained in the clustered markets case; robust standard errors approximate the true standard errors well. On the other hand, non-robust standard errors for $\widehat{\alpha}$ and $\widehat{\beta}_0$ underestimate the true standard errors, leading to under-coverage of the nominal 95% confidence intervals.

Table 3: Standard errors in markets on a line and a lattice

	Line				Lattice			
Parameter	α	β_0	β_1	γ	α	β_0	β_1	γ
Average $\widehat{\theta}$	1.009	1.010	0.999	0.501	1.008	1.009	1.000	0.501
Std. dev. $\widehat{\theta}$	0.126	0.135	0.060	0.043	0.102	0.110	0.060	0.043
Robust s.e.	0.125	0.136	0.060	0.043	0.100	0.110	0.060	0.043
95%-coverage	0.949	0.952	0.949	0.949	0.945	0.949	0.949	0.949
Non-robust s.e.	0.103	0.113	0.060	0.043	0.085	0.096	0.060	0.043
95%-coverage	0.891	0.900	0.949	0.948	0.900	0.910	0.948	0.949

Notes: Each experiment is based on 2,000 simulation draws. Average $\widehat{\theta}$ is the sample mean of the simulated estimates, and Std. dev. $\widehat{\theta}$ is the standard deviation of the estimates. Non-robust s.e. is the mean of the estimated standard errors that are only robust against heteroscedasticity. Robust s.e. is the mean of the estimated standard errors that are robust against correlation between g_{jt}s. 95%-coverage calculates the actual coverage of a nominal 95% confidence interval formed by using the respective standard error.

5 Conclusion

In this paper, I address the previously overlooked issue of correlation in the sample moments across markets in demand estimation, induced by the essential endogeneity of Hausman instruments. I show that common methods of calculating standard errors can be inaccurate by not accounting for the correlation, and propose practical and easily implementable remedies in leading cases of Hausman instruments. When using a region-based Hausman instrument, one needs to use a standard error clustered at the region level. When using an adjacency-based Hausman instrument, one needs to use an autocorrelation-robust standard error or a standard error that accounts for cross-sectional dependence. The key takeaway is that one needs to examine, as demonstrated in the paper, the correlation structure that emerges from the researcher's choice of the Hausman instrument, and use an appropriate variance estimator.

In the clustered markets context, Monte Carlo simulations indicate that clustered standard errors are accurate with a large number of clusters. For a small number of clusters, the conventional standard errors, despite their inaccuracies, might be preferable. The true variance of the estimator tends to be smaller when the actual data generating process is that the actual regions, within which the cost shifters are common, are moderately sized. Based on these findings, it is recommended to find moderately sized regions within which some cost shifters are believed to be common, and then apply clustering to accurately estimate the true standard error.

A An example of nonzero correlation

For a simple analytic illustration that g_{jt} can be correlated across markets, suppose each market is a monopoly with the demand function $q(p) = 2 - p + \xi_t$ and the marginal cost $c_t = w_t$. To incorporate this into the discrete choice model framework, we may set $u_{i1t}(p) = 2 - p + \xi_t - \epsilon_{i1t}$ and $u_{i0t}(p) = 0$ with a constant market size 1, where ϵ_{i1t} is uniformly distributed on $[0, 1]$. Assume that ξ_t and ξ_t' are independently drawn from $\{-1, 1\}$ with equal probability, independent from $w = (w_t, w_t')$. The distribution of w may be left unspecified, but assume that the support of w_t and w_t' lies in $[0, 1]$ for convenience. Then the equilibrium price is determined by $p_t = 1 + (\xi_t + w_t)/2$, and we have $\mathbb{E}[g_t g_{t'}] = \mathbb{E}[p_{t'}\xi_t p_t \xi_{t'}] = 1/4 > 0$.

B More results on misspecified regions

The following figures plot the actual coverage probability of the nominal 95% confidence intervals (CIs) based on the robust- and non-robust standard errors. In each plot, the x-axis is the scope of the Hausman instrument, i.e., the size of each cluster used in the clustered standard error. Vertical dashed lines are positioned to denote the size of the actual region, i.e., the region in which the cost shifter is constant across markets.

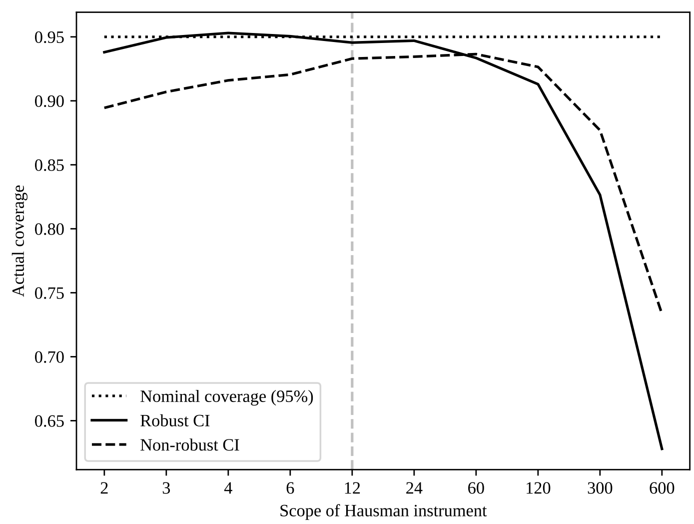


Figure 4: Coverage of CIs in clustered markets (misspecified regions, true = 12)

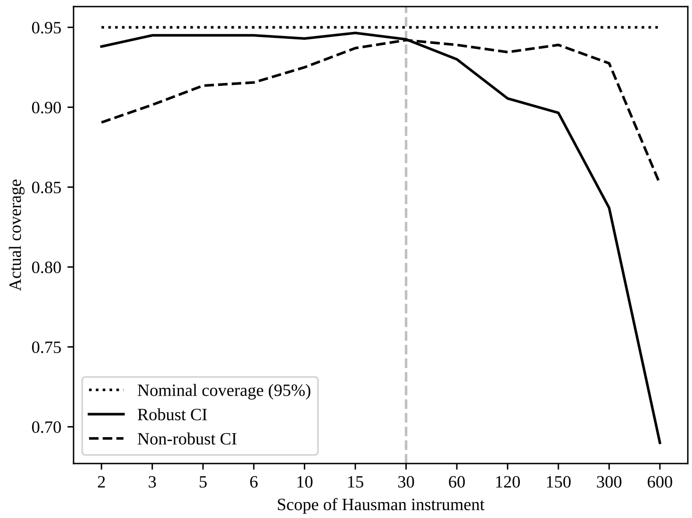


Figure 5: Coverage of CIs in clustered markets (misspecified regions, true = 30)

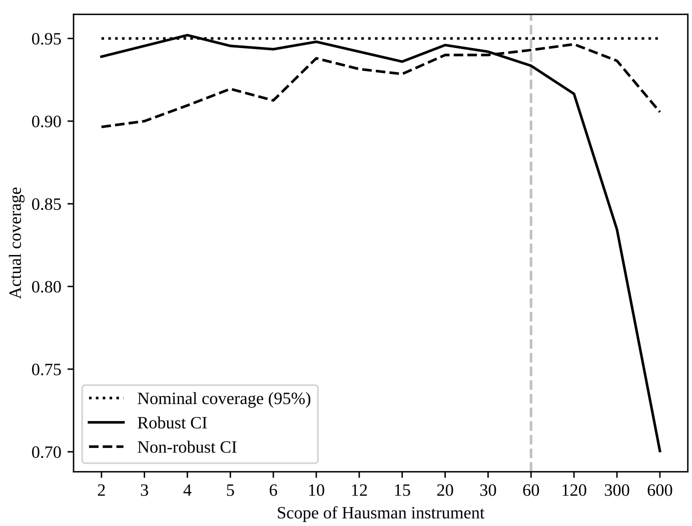


Figure 6: Coverage of CIs in clustered markets (misspecified regions, true = 60)

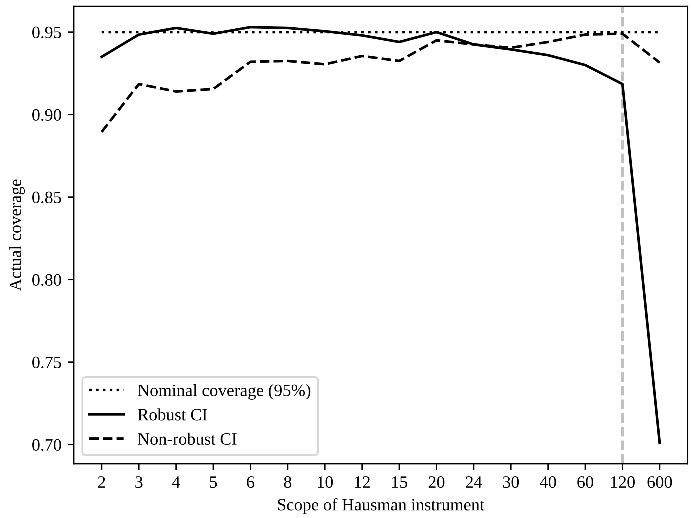


Figure 7: Coverage of CIs in clustered markets (misspecified regions, true = 120)

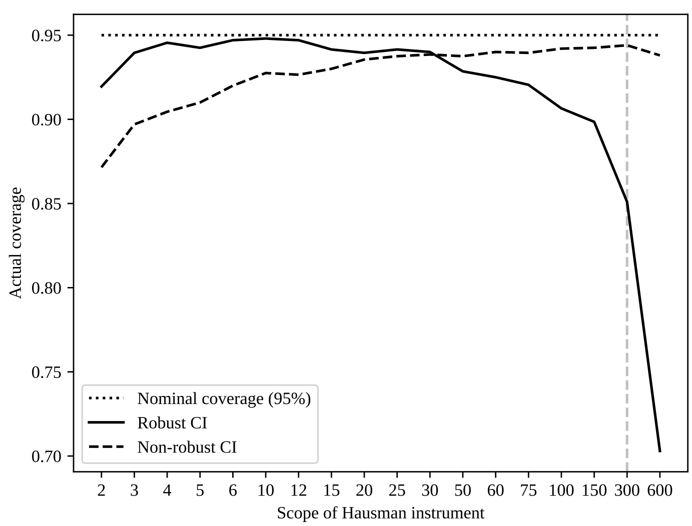


Figure 8: Coverage of CIs in clustered markets (misspecified regions, true = 300)

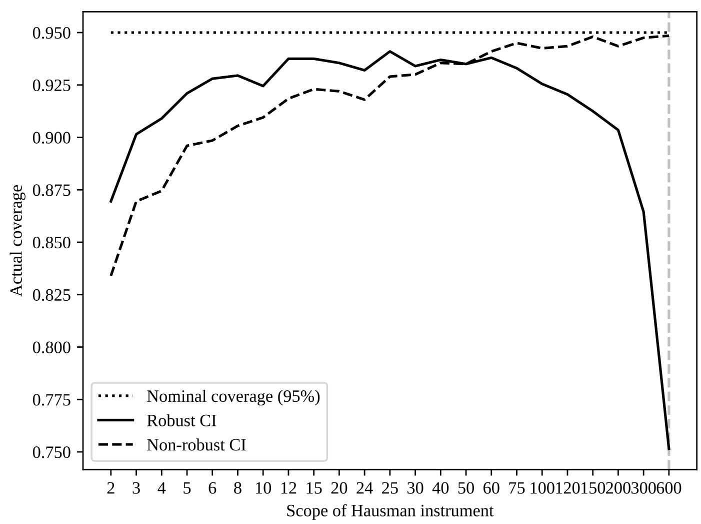


Figure 9: Coverage of CIs in clustered markets (misspecified regions, true = 600)

References

Berry, S., A. Gandhi, and P. Haile (2013). Connected substitutes and invertibility of demand. *Econometrica 81*(5), 2087–2111.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

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

Berry, S. T. and P. A. Haile (2014). Identification in differentiated products markets using market level data. *Econometrica 82*(5), 1749–1797.

Berry, S. T. and P. A. Haile (2021). Foundations of demand estimation. In *Handbook of Industrial Organization*, Volume 4, pp. 1–62. Elsevier.

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

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics 92*(1), 1–45.

Conlon, C. and J. Gortmaker (2020). Best practices for differentiated products demand estimation with PyBLP. *The RAND Journal of Economics 51*(4), 1108–1161.

Gandhi, A. and J.-F. Houde (2019). Measuring substitution patterns in differentiated-products industries. *NBER Working Paper* (w26375).

Guevara, C. A. and M. Ben-Akiva (2006). Endogeneity in residential location choice models. *Transportation Research Record 1977*(1), 60–66.

Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In *The Economics of New Goods*, pp. 207–248. University of Chicago Press.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, 395–421.

Newey, W. K. and K. D. West (1987, 05). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica (1986-1998) 55*(3), 703.

Ogaki, M. (1993). Generalized method of moments: Econometric applications. In *Handbook of Statistics*, Volume 11, pp. 455–488. Elsevier.