Competition among Spatially Differentiated Firms: An Empirical Model with an Application to Cement*

Nathan H. Miller
Department of Justice

Matthew Osborne
Department of Commerce

September 2010

Abstract

The theoretical literature of industrial organization shows that the distances between consumers and firms have first-order implications for competitive outcomes whenever transportation costs are large. To assess these effects empirically, we develop an estimator for models of spatial differentiation and spacial price discrimination that recovers the underlying structural parameters using only aggregate data. We provide conditions under which the estimates are consistent and asymptotically normal. We apply the estimator to the portland cement industry. The estimation fits, both in-sample and out-of-sample, demonstrate that the framework explains well the salient features of competition. We estimate transportation costs to be $0.30 per tonne-mile and show that these costs constrain shipping distances and create localized market power. To demonstrate policy-relevance, we conduct counter-factual simulations that quantify competitive harm from a hypothetical merger. We map the distribution of harm over geographic space and identify the divestiture that best mitigates harm.

Keywords: transportation costs; spatial differentiation; price discrimination; cement
JEL classification: C51; L11; L40; L61

*Miller: nathan.miller@usdoj.gov. Osborne: matthew.osborne@bea.gov. We thank Masakazu Ishihara, Robert Johnson, Scott Kominers, Ashley Langer, Russell Pittman, Chuck Romeo, Jim Schmitz, Adam Shapiro, Charles Taragin, Raphael Thomadsen, Glen Weyl and seminar participants at the Bureau of Economic Analysis, the Federal Trade Commission, Georgetown University, the University of British Columbia – Sauder School of Business, and the U.S. Department of Justice for valuable comments. Sarolta Lee, Parker Sheppard, Gian Wrobel, and Vera Zavin provided research assistance. The views expressed herein are entirely those of the authors and should not be purported to reflect those of the U.S. Department of Justice or attributed to the Bureau of Economic Analysis.

1 Introduction

In many industries, firms are geographically differentiated and transportation is costly. Yet few empirical studies estimate structural models of spatial differentiation. We attribute this dearth of research to a simple *data availability problem*: the most straight-forward way to identify the degree of spatial differentiation – or, equivalently, the magnitude of transportation costs – is to measure how firms' market shares differ between nearby and distant consumers. But this requires data on the geographic distributions of the market shares. These data are difficult to attain and, indeed, we are unaware of any study that exploits variation in market shares over geographic space.

The data availability problem is only exacerbated for industries characterized by spatial price discrimination because it becomes necessary to account for the geographic distributions of the prices, as well. While three recent studies apply econometric techniques to sidestep the data availability problem in non-discriminatory settings (Thomadsen (2005), Davis (2006), McManus (2009)),[1] to our knowledge no previous work estimates structural parameters in the face of spatial price discrimination – despite the oft-cited result of Greenhut, Greenhut, and Li (1980) that 32 percent of surveyed firms employ some form of spatial price discrimination.[2]

This sparse empirical literature is in contrast to a storied theoretical literature (e.g., Hotelling (1929), d'Aspremont, Gabszewicz, and Thisse (1979), Salop (1979), Thisse and Vives (1988), Economides (1989), Vogel (2008)). Indeed, it is well understood that spatial differentiation has first-order implications for competitive outcomes whenever transportation costs are large. But how large is large? Without a viable estimation strategy, the insights of the theoretical literature can be difficult to apply in specific, real-world settings – the settings that, presumably, matter most.

In this paper, we develop a new estimation strategy for economic models of spatial differentiation and spatial price discrimination. The estimator exploits variation in aggregated data that are commonly available to the econometrician – such as average prices, total production, or regional consumption. The estimator does not suffer from the data availability problem because data on the spatial distributions of the market shares and prices are not needed. We provide conditions under which the estimates are consistent and asymptotically

[1]Pinske, Slade, and Brett (2002) develop a useful reduced-form estimator that can be applied to evaluate the extent to which competition is localized in a non-discriminatory setting. The estimator does not recover the underlying structural parameters of the model, including the transportation cost, and therefore does not enable counter-factual policy experimentation. See Section 2 for details.

[2]The Greenhut, Greenhut, and Li (1980) sample is small and not clearly representative. More systematic efforts to identify those industries that employ spatial price discrimination have not been made.

normal. We also conduct an empirical application and demonstrate that (1) estimation is feasible for real-world data; (2) the salient features of competition are explainable with only limited data; and (3) the results enable powerful new counter-factual policy experiments.

The central insight is that one can solve the data availability problem by relying on numerical approximations to equilibrium. That is, one can *compute* the geographic distributions of markets shares and prices that characterize the equilibrium of the economic model, for any candidate parameter vector. These distributions can be aggregated to match the data, and one can search for the parameter vector that minimizes the "distance" between the predictions and the data. We find it intuitive to think of the procedure as having an outer loop and an inner loop. In the outer loop, an objective function is minimized over the parameter space; in the inner loop, equilibrium is computed numerically for each candidate parameter vector. This structure makes our estimator broadly analogous to other estimators developed for discrete static games (e.g., Bajari, Hong, and Ryan (2008)), non-strategic dynamic games (e.g., Rust (1987)) and certain strategic dynamic games (e.g., Goettler and Gordon (2009), and Gallant, Hong, and Khwaja (2010)), in the sense that each requires the repeated computation of equilibrium.[3]

The reliance on numerical approximations to equilibrium also sidesteps any complications associated with measurement error in the data. The existing methods for estimating models of spatial differentiation (e.g., Thomadsen (2005), Davis (2006), McManus (2009)) use firm-level price data to construct the right-hand-side of the regression equation. If measurement error is present then it can propagate nonlinearly and create biases in the regression coefficients. Such concerns do not arise in our approach provided that the numerical approximations to equilibrium are sufficiently precise.

Importantly, the power of modern computers makes the estimation strategy we introduce feasible if one is willing to make convenient functional form assumptions on demand and marginal costs. These assumptions, though potentially restrictive, nonetheless permit the estimation of economic models that are more realistic than the standard market delineation model, which "solves" the data availability problem using the dual assumptions that (1) transportation costs are large enough to preclude competition across market boundaries, and (2) transportation costs are small enough that spatial differentiation is negligible within

[3]The key empirical relationships that drive parameter estimates within our framework can be quite transparent despite the complicated nonlinear relationships involved. In our application, the transportation cost estimate is driven by differences between consumption and production within specific geographic regions. Suppose, for instance, that one observes that total consumption is greater than total production in one region but less than total production in another. This implies that inter-regional trade flows exist, and a parameter can be selected that rationalizes these trade flows within the structure of the model.

markets.[4] These assumptions preclude inference regarding spatial differentiation because the transportation cost cannot be estimated structurally. Further, markets tend to be delineated based on political borders of questionable economic significance such as state or county lines. Yet this approach has been employed routinely to study of industries characterized by high transportation costs, including ready-mix concrete (e.g., Syverson (2004), Syverson and Hortaçsu (2007), Collard-Wexler (2009)), portland cement (e.g., Salvo (2008), Ryan (2009)), and paper (e.g., Pesendorfer (2003)).[5]

In the empirical application, we examine the portland cement industry in the U.S. Southwest over the period 1983-2003. The available data include average prices, production, and consumption, each at the regional level (e.g., we observe total consumption separately for northern California, southern California, Arizona and Nevada). We find that the estimation procedure produces impressive in-sample and out-of-sample fits despite parsimonious demand and marginal cost specifications. For instance, the model predictions explain 93 percent of the variation in regional consumption, 94 percent of the variation in regional production, and 82 percent of the variation in regional prices. The model predictions also explain 98 percent of the variation in cross-region shipments, even though we withhold the bulk of these data from estimation. The quality of these fits is underscored by the rich time-series variation in these data due to macro-economic fluctuations.

Pushing the results further, we estimate that consumers pay roughly \$0.30 per tonne-mile, given diesel prices at the 2000 level.[6,7] This translates into an average transportation cost of \$24.61 per metric tonne, or 22 percent of total consumer expenditure over the sample period. We find that cement travels 92 miles on average and that firms set higher prices to nearby consumers and consumers without viable alternatives, yet maintain greater market shares among these consumers. (These patterns are as expected since consumers pay the costs of transportation in the cement industry.) Thus, the results help quantify the extent to which cement firms exercise localized market power over their consumers.

[4]Syverson (2004) discusses how this tension can compel researchers to seek compromise between markets that are "too small" and markets that are "too large". It is sometimes argued that markets that are too large overstate the intensity of competition while markets that are too small understate competition.

[5]These valuable contributions focus on wide range of topics, including the competitive impacts of horizontal and vertical mergers, heterogeneity in plant productivity and its implications for competition, the inference of market power, and dynamic investment decisions.

[6]The 1974 edition of the Minerals Yearbook, an annual publication of the U.S. Geological Survey, identifies a per-tonne transportation cost of \$0.35 (when adjusted for inflation). Subsequent editions do not provide the magnitude of transportation costs.

[7]The model identifies consumer willingness-to-pay for proximity to the plant, which incorporates transportation costs as well as any other distance-related costs (e.g., reduced reliability). We refer to this willingness-to-pay as the transportation cost, although the concepts may not be precisely equivalent.

We finish by evaluating the effects of a hypothetical merger between two multi-plant firms. The value of our approach is easily discernable in such a policy application because we are able to map the distribution of harm over the U.S. Southwest and identify which divestiture(s) would most effectively mitigate or eliminate harm. In the exercise, we find that the merger reduces consumer surplus by $1.4 million and that harm is concentrated in counties surrounding Los Angeles and Phoenix. We consider each of the possible single-plant divestitures and find that the most powerful reduces harm by 56 percent.[8]

Finally, we note that spatial considerations create interesting dynamics. Firms may select their locations to secure a base of profitable customers, provide separation from efficient competitors, or deter nearby entry. We abstract from these considerations in our application and instead assume that firm location is pre-determined and exogenous.[9] Nonetheless, our estimation strategy could be applied to define stage-game payoffs in certain dynamic strategic games (e.g., those of Seim (2006) and Ryan (2009)) and our present work can be interpreted as a useful first step in the examination of dynamic spatial competition.

The paper proceeds as follows. We first review the relevant empirical literature in Section 2 and clarify our contribution relative to this literature. We then formalize a model of spatial price discrimination in Section 3. We develop the estimation strategy and derive the asymptotic properties of the estimator in Section 4. We then turn to the empirical application. Section 5 develops the institutional details of the portland cement industry and describes the available data. In Section 6, we specify the model, derive the objective function, conduct an identification test, discuss multiple equilibria, and plot the identifying variation in the data. We present and evaluate the results in Section 7, and then conclude.

2 Review of the Empirical Literature

The closest antecedents of our work are Thomadsen (2005), Davis (2006) and McManus (2009), each of which estimates a model of spatial differentiation in a non-discriminatory setting (fast-food restaurants, movie theaters, and coffee shops, respectively). Together, the papers demonstrate that the data availability problem is solvable provided that there is no spatial price discrimination and that the econometrician has two of the following: data on firm-level prices, data on firm-level market shares, and an equilibrium condition derived from some underlying economic model.

Thomadsen (2005) combines firm-level prices and an equilibrium condition to solve for

[8]We do not attempt to characterize the appropriate course of action for an antitrust authority.

[9]The assumption may be justifiable given that we observe only two plant closures and one plant entry.

the geographic distribution of market shares at each candidate parameter vector. (These market shares have convenient analytical solutions given the assumed logit demand function.) The estimation procedure then selects the parameters that bring the implied equilibrium firm-level prices close to the data. By contrast, Davis (2006) and McManus (2009) exploit variation in firm-level prices and sales. They derive predicted sales in a number of sub-markets for each candidate parameter vector, given prices and an observed geographic distribution of consumers, and aggregate these predictions to construct predicted firm-level sales. The estimation procedures then selects the parameters that bring the predicted firm-level sales close to the data. Though Davis and McManus focus solely on the demand parameters, their methods could be used to impute the usual supply constructs (e.g., marginal costs) via an equilibrium condition.

Our insight is that estimation is feasible with *only* the equilibrium condition: neither firm-level prices nor firm-level market shares must be observed by the econometrician. The more flexible data requirements may enable research on industries for which a full complement of firm-level data is unavailable. This may prove particularly valuable to antitrust authorities because it can be difficult to obtain all the relevant firm-level prices in merger investigations when third parties do not cooperate. The greater reliance on the equilibrium condition also has two other advantages: First, it makes feasible the estimation of spatial price discrimination models because one need not observe the geographic distribution of prices in the data. Second, the consistency of the estimates is robust to the presence of measurement error in the data. Such measurement error can propagate nonlinearly in frameworks of Thomadsen (2005), Davis (2006) and McManus (2009).

Finally, we discuss the contribution of Pinske, Slade, and Brett (2002), which develops a reduced-form estimator for non-discriminatory spatial differentiation models. The authors regress prices on competitors' prices and controls – the coefficients on competitors' prices are assumed to be nonparametric functions of the distances between firms and dictate how spatial differentiation softens price competition. The model can be usefully applied to evaluate the extent to which competition is localized. Further, the estimator easily accommodates unobserved firm heterogeneity if instruments are available.[10] The estimator does not recover the structural parameters of the model, however, because the estimated price coefficients are complicated combinations of the underlying parameters. As such, the results do not enable counter-factual policy experiments.

[10]In principle, it is possible to incorporate unobserved firm heterogeneity in our model using simulation techniques. One would have to compute equilibrium for each simulation draw which would increase the computer resources needed for estimation.

3 The Model of Price Competition

3.1 The geographic space

We define the relevant geographic *space* to be a compact, connected set $\mathbb{C}$ in the Euclidean space $\mathbb{R}^2$. We take as given that J plants compete in the space, and assume that each plant is endowed with a fixed location defined by the geographic coordinates $\{z_1, z_2, \ldots, z_J\}$, where $z_j \in \mathbb{C}$. We further take as given that a continuum of consumers spans the space, and assume that each consumer has unit demand and a fixed location $w \in \mathbb{C}$. The absolute measure $\phi(w)$ characterizes the geographic distribution of consumers and we define $M = \int_{\mathbb{C}} \phi(w)dw$ to be the potential demand of the space. We denote the distance between any two points in the geographic space, say a and b, as the Euclidean distance $\|a - b\|$.

Without loss of generality, we partition the geographic space into N distinct geographic consumer *areas*, such that each area $\mathbb{C}_n \in \mathbb{C}$ is itself a connected set in $\mathbb{R}^2$. We conduct the partition such that $\mathbb{C}_1 \bigcup \mathbb{C}_2 \bigcup \cdots \bigcup \mathbb{C}_N = \mathbb{C}$, and $\mathbb{C}_r \bigcap \mathbb{C}_s = \emptyset$ for any $r \neq s$. The function $G : \mathbb{R}^2 \rightarrow \{1, ..., N\}$ maps consumers to areas and the potential demand of each area is given by $M_n = \int_{\mathbb{C}_n} \phi(w)dw$. The partition does not circumscribe competition because any plant can ship to any consumer, but is nonetheless central to the economics of the model and we provide explicit motivation in Section 3.4.

3.2 Supply and demand

We assume that F firms compete in prices. Each firm operates some subset $\mathbb{J}_f$ of the J plants and can ship from any plant $j \in \mathbb{J}_f$ to any consumer. We assume that firms are able to employ imperfect spatial price discrimination by setting different mill prices to different areas. Note that in our model, the mill price does not include the transport cost: a consumer's total payment for the product will be the mill price plus the door-to-door transportation cost. Let the price vector $\boldsymbol{p}_n \in \mathbb{R}^J$ characterize the mill prices available to consumers in area $\mathbb{C}_n$, and let the vector $\boldsymbol{z}$ include the plant locations. Then firms maximize variable profits:

$$\pi_f = \underbrace{\sum_{j \in \mathbb{J}_f} \sum_n p_{jn} q_{jn}(\boldsymbol{p_n}; \boldsymbol{z}, \boldsymbol{x}_{jn}, \boldsymbol{\theta}_0)}_{\text{variable revenues}} - \underbrace{\sum_{j \in \mathbb{J}_f} \int_0^{Q_j(\boldsymbol{p}; \boldsymbol{z}, \boldsymbol{X}_j, \boldsymbol{\theta_0}))} c(Q; \boldsymbol{w}_j, \boldsymbol{\theta}_0) dQ}_{\text{variable costs}}, \tag{1}$$

where p_{jn} is the mill price of plant j to consumers in area $\mathbb{C}_n$, $q_{jn}(\boldsymbol{p_n}; \boldsymbol{z}, \boldsymbol{x}_{jn}, \boldsymbol{\theta}_0)$ is the quantity produced by plant j and sold to consumers in area $\mathbb{C}_n$, $Q_j(\boldsymbol{p}; \boldsymbol{z}, \boldsymbol{X}_j, \boldsymbol{\theta}_0)$ is the total quantity produced by plant j, and $c(Q; \boldsymbol{w}_j, \boldsymbol{\theta}_0)$ is a convex and differential marginal cost function of known form. The vectors $\boldsymbol{x}_{jn}$ and $\boldsymbol{w}_j$ include demand and cost shifters, respectively, and the matrix $\boldsymbol{X}_j$ stacks the relevant $\boldsymbol{x}_{jn}$ vectors. Finally, $\boldsymbol{\theta}_0$ is a K-dimensional parameter vector.

We model consumer behavior using a conventional discrete-choice demand system. Each consumer observes the plant locations and the available mill prices, and either purchases from one of the J plants or foregoes a purchase altogether (i.e., selects the outside good). The indirect utility that consumer i receives from plant j is:

$$u_{ij} = \beta^c + \beta^p p_{nj} + \beta^d \|w - z_j\| + \boldsymbol{x}'_{jn}\boldsymbol{\beta}^x + \nu_{ij}, \tag{2}$$

where ν_{ij} is an idiosyncratic preference shock that is observed to the consumer and uncorrelated with distance and prices. We provide motivation for the preference shock to Section 3.4. Following standard practice, we normalize the mean utility of the outside option to zero. Finally, $(\beta^c, \beta^p, \beta^d, \boldsymbol{\beta}^x) \in \boldsymbol{\theta}_0$ are the demand parameters and the ratio β^d/β^p represents the unit transportation cost incurred by consumers.

We assume that consumers select the plant that supplies the highest utility. Within an area, this assumption defines the set of consumer characteristics $(w \in \mathbb{C}_n, \boldsymbol{\nu}_i)$ that lead to the selection of plant j, and we denote this set

$$A_{jn}(\boldsymbol{p}_n; \boldsymbol{z}, \boldsymbol{x}_{jn}, \boldsymbol{\theta}_0) = \{(w \in \mathbb{C}_n, \boldsymbol{\nu}_i) | u_{ij} \geq u_{ik} \ \forall\ k = 0, 1, \ldots J\},$$

where $\boldsymbol{z} = (z_1, z_2, \ldots, z_J)$. If ties occur with zero probability then the quantity produced by plant j and consumed in area $\mathbb{C}_n$ is given by:

$$q_{jn}(\boldsymbol{p_n}; \boldsymbol{z}, \boldsymbol{x}_{jn}, \boldsymbol{\theta}_0) = M_n \int_{A_{jn}} \partial P(w, \nu), \tag{3}$$

where $P(\cdot)$ denotes a population distribution function. We place two normalcy conditions on demand, namely that $q_{jn}(\boldsymbol{p_n}; \boldsymbol{z}, \boldsymbol{x}_{jn}, \boldsymbol{\theta}_0)$ is twice continuously differentiable and also downward sloping in p_{jn} (i.e., $\beta^p < 0$).

For clarity, we sketch one possible geographic space in Figure 1. The dashed lines delineate three consumer areas, $\mathbb{C}_1$, $\mathbb{C}_2$, and $\mathbb{C}_3$. Two plants operate in the space and are characterized by the locations z_1 and z_2. A distribution of consumers span the space, and both plants compete for each consumer. The plants imperfectly price discriminate by setting different prices for consumers in each area. Thus, there are six prices in the space, which

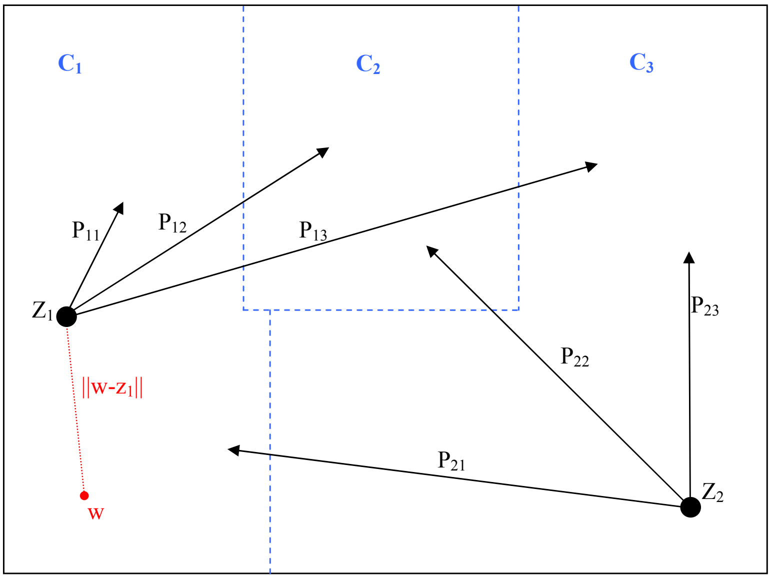

Figure 1: A Geographic Space.

we represent with the arrows labeled $\{p_{11}, p_{12}, p_{13}, p_{21}, p_{22}, p_{23}\}$. Finally, we plot the location of a single consumer characterized by location w. The dashed line labeled $\|w - z_1\|$ is the Euclidean distance between the consumer and the first plant.

3.3 Equilibrium

The variable profit function of equation (1) yields first-order conditions that characterize each firm's best-response vector of mill prices. To ease notation, we define a set ζ that includes the plant locations as well as the demand and cost shifters. The first-order conditions are:

$$q_{jn}(\boldsymbol{p_n}; \zeta, \boldsymbol{\theta}_0) + \sum_n \sum_{k \in \mathbb{J}_f} (p_{kn} - c_k(Q_k(\boldsymbol{p}; \zeta, \boldsymbol{\theta}_0); \zeta, \boldsymbol{\theta}_0)) \frac{\partial q_{kn}(\boldsymbol{p_n}; \zeta, \boldsymbol{\theta}_0)}{\partial p_{jn}} = 0. \tag{4}$$

There are $J \times N$ first-order conditions. For notational convenience, we define the block-diagonal matrix $\boldsymbol{\Omega}(\boldsymbol{p}; \zeta, \boldsymbol{\theta}_0)$ as the combination of $n = 1, \ldots, N$ sub-matrices, each of dimension $J \times J$. The elements of the sub-matrices are defined as follows:

$$\Omega^n_{jk}(\boldsymbol{p_n}; \zeta, \boldsymbol{\theta}_0) = \begin{cases} \frac{\partial q_{jn}(\boldsymbol{p_n}; \zeta, \boldsymbol{\theta}_0)}{\partial p_{kn}} & \text{if } j \text{ and } k \text{ have the same owner} \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

The elements of each sub-matrix characterize substitution patterns within area $\mathbb{C}_n$, and $\boldsymbol{\Omega}$ has a block diagonal structure because $q_{jn}(\boldsymbol{p_n}; \zeta, \boldsymbol{\theta}_0)$ is free of $\boldsymbol{p}_{-n}$. Thus, the construction of $\boldsymbol{\Omega}$ builds on the premises that (1) consumers in each area $\mathbb{C}_n$ select among all J plants, and (2) demand in area $\mathbb{C}_n$ is unaffected by mill prices in area $\mathbb{C}_m$ for $n \neq m$.

We now rearrange and stack the first-order conditions:

$$\boldsymbol{f}(\boldsymbol{p};\zeta,\boldsymbol{\theta}_0) \equiv \boldsymbol{p} - \boldsymbol{c}(\boldsymbol{Q}(\boldsymbol{p};\zeta,\boldsymbol{\theta}_0);\zeta,\boldsymbol{\theta}_0) + \boldsymbol{\Omega}^{-1}(\boldsymbol{p};\zeta,\boldsymbol{\theta}_0)\boldsymbol{q}(\boldsymbol{p};\zeta,\boldsymbol{\theta}_0) = \mathbf{0}. \tag{6}$$

A vector of prices that solves this system of equations is a spatial Bertrand-Nash equilibrium. We define a mapping $\boldsymbol{H}(\boldsymbol{\theta}_0;\zeta) : \mathbb{R}^K \to \mathbb{R}^{JN}$ that matches the parameters of the model to spatial Bertrand-Nash equilibrium given the exogenous data. Formally, the mapping is defined by the equivalence $\boldsymbol{f}(\boldsymbol{H}(\boldsymbol{\theta}_0;\zeta);\zeta,\boldsymbol{\theta}_0) \equiv \mathbf{0}$.

3.4 Discussion

We offer three comments to help build intuition on the economics of the model. First, spatial price discrimination is at the core of the firm's pricing problem: firms charge higher mill prices to nearby consumers and to consumers for whom the firm's competitors are more distant. However, aside from price discrimination, the firm's pricing problem follows standard intuition. A firm that contemplates a higher mill price from one of its plants to a given area must evaluate (1) the tradeoff between lost sales to marginal consumers and greater revenue from inframarginal consumers; and (2) whether the firm would recapture lost sales with its other plants. If marginal costs are not constant, then the firm must also evaluate how the lost sales would affect the plant's competitiveness in other areas.

Second, the areas $\mathbb{C}_1, \mathbb{C}_2, \ldots \mathbb{C}_N$ are best interpreted as determining the extent which firms engage in spatial price discrimination. Finer partitions of the geographic space correlate with more sophisticated discrimination, and if only a single area exists (i.e., $N = 1$) then firms do not discriminate. *The areas have no economic significance aside from these implications for spatial price discrimination.* Since every plant competes in every area, the partition of the geographic space into distinct areas does not artificially limit competition and is not analogous to a "market delineation" assumption under which plants compete only within prescribed geographic boundaries.

Finally, the indirect utility specification of equation (2) implies that plants are differentiated by both location and idiosyncratic preferences shocks. In the special case of degenerate preference shock distribution, the model collapses to a "pure characteristics model" along the lines of the original Hotelling (1929) formulation. Although the estimation strategy we outline below accommodates the pure characteristics model on a theoretical level, we suspect that implementation would be frustrated by complications that we discuss in our application (see Section 6.1.2). In the more general case, the idiosyncratic preference shocks can be motivated as capturing various plant- and consumer-level heterogeneity that, due to costly bargaining or other reasons, is not reflected in mill prices.

4 Estimation

The model generates a rich set of predictions on equilibrium prices, production, and shipments within the geographic space $\mathbb{C}$. Yet the parameters of the model can be recovered using relatively coarse data. In this section, we introduce a novel GMM estimation procedure that exploits variation in $t = 1, 2, ...T$ time-series observations on aggregated prices (e.g., observations on average mill prices). We show that the GMM estimator is consistent and asymptotically normal, given assumptions on the existence and uniqueness of equilibrium. We then demonstrate that the estimator can be extended in a straight-forward manner to exploit variation in other endogenous data, such as observations on production and consumption. The flexibility of these data requirements makes the estimator widely applicable to economic settings characterized by spatial differentiation.

4.1 GMM estimation

We first clarify the level of detail on consumer locations, $w \in \mathbb{C}$, needed to support estimation. In many instances, precise consumer locations may be unavailable or too costly to discover, so that the direct application of equation (2) is infeasible. We make the following assumption on the exogenous spatial data available to the econometrician:

Assumption A1: *The econometrician observes the mean distance between plant j and the consumers in area $\mathbb{C}_n$, for all j and n.*

We denote the mean distance between plant j and the consumers in area $\mathbb{C}_n$ as d_{jnt}. Under A1, we can rewrite the indirect utility equation as follows:

$$u_{ijt} = \beta^c + \beta^p p_{njt} + \beta^d d_{njt} + \boldsymbol{x}'_{jnt}\boldsymbol{\beta}^x + \nu^*_{ijt}, \tag{7}$$

where ν^*_{ijt} is a composite error term that includes the idiosyncratic preference shock and the consumer-specific deviation from mean distance. Formally, $\nu^*_{ijt} = \nu_{ijt} + \beta^d(\|w - z_{jt}\| - d_{njt})$. The composite error term is orthogonal to price and mean distance, given the assumptions already placed on the model. As long as the distributions of ν^*_{ijt} are known, or reasonable approximations can be made, compute demand can be computed given the relevant prices and the mean distances between plants and areas. We formalize this in Assumption A2.

Assumption A2: *The econometrician knows the distributions of ν^*_{ijt}.*

Integrating over this distribution yields an equilibrium mapping $\boldsymbol{H}(\boldsymbol{\theta}_0; \zeta_t)$ that depends on mean distances and the demand and cost shifters. We assume that the price data are generated by the following process:

$$\boldsymbol{p}^{\epsilon}_t = \boldsymbol{H}(\boldsymbol{\theta}_0; \zeta_t) + \boldsymbol{\epsilon}_t, \tag{8}$$

where $\boldsymbol{p}^{\epsilon}_t$ is a vector of length JN, and $\boldsymbol{\epsilon}_t$ is a vector of unobserved sampling errors. The vector $\boldsymbol{p}^{\epsilon}_t$ will not typically be observed in the data, and instead we assume that the econometrician observes prices on a more aggregate level. Examples could include the average price charged by each plant, the average price charged by various combinations of plants (e.g, firm-specific average prices), or the average price paid by consumers within certain geographic regions (perhaps spanning multiple areas). We define a function $\boldsymbol{S} : \mathbb{R}^{JN} \rightarrow \mathbb{R}^L$ that maps the plant-area prices to the aggregate prices. Denoting the vector of observed aggregate prices as $\boldsymbol{p}^d_t$, we have:

$$\boldsymbol{p}^d_t \equiv \boldsymbol{S}(\boldsymbol{p}^{\epsilon}_t) = \boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta}_0; \zeta_t) + \boldsymbol{\epsilon}_t). \tag{9}$$

We make the following assumption about the aggregation process:

Assumption A3: *The function $\boldsymbol{S}$ is linear.*

Under A3, the aggregated sampling errors are additively separable, i.e., $\boldsymbol{p}^d_t = \boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta}_0; \zeta_t)) + \boldsymbol{S}(\boldsymbol{\epsilon}_t)$. The assumption holds for most data that econometricians encounter in practice; aggregation processes based on averages, for example, are clearly consistent with the assumption.

We need to make one more assumption to construct our GMM estimator:

Assumption A4: *There exist instruments such that*

$$E[\boldsymbol{p}^d_t - \boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta}_0; \zeta_t)) | \, \boldsymbol{Z}_t] = \boldsymbol{0},$$

where $\boldsymbol{Z}_t = [\boldsymbol{z}^1_t \; \boldsymbol{z}^2_t \ldots \boldsymbol{z}^M_t]'$ is an $M \times L$ matrix that combines the $M \geq K$ instrument vectors.

A4 generates the population moment equations $E[\boldsymbol{Z}_t(\boldsymbol{p}^d_t - \boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta}_0; \zeta_t)))] = \boldsymbol{0}$, and these moment equations can support GMM estimation. In many situations, however, it may be reasonable to further assume that the sampling error is independent of the "right-hand-side" data ζ_t. This simplifies the construction of the estimator, and we impose the additional assumption here:

Assumption A4′: *The sampling error is mean zero conditional on* ζ_t*:*

$$E[\boldsymbol{p}_t^d - \boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta}_0; \zeta_t))|\, \zeta_t] = \mathbf{0}.$$

A4′ enables estimation with multiple equation nonlinear least squares, which is equivalent to GMM with the optimal instruments

$$\boldsymbol{Z}_t = -\frac{\partial \boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta}_0; \zeta_t))}{\partial \boldsymbol{\theta}_0} \boldsymbol{\Lambda}_0(\boldsymbol{\theta}_0)^{-1}, \tag{10}$$

where $\boldsymbol{\Lambda}_0(\boldsymbol{\theta}_0) \equiv E[\boldsymbol{S}(\boldsymbol{\epsilon}_t)|\zeta_t]E[\boldsymbol{S}(\boldsymbol{\epsilon}_t)|\zeta_t]'$ is the variance matrix of the aggregated error terms. Thus, the sample moment equations that correspond to A4′ are

$$\frac{1}{T}\sum_{t=1}^{T} -\frac{\partial \boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta}; \zeta_t))}{\partial \boldsymbol{\theta}} \boldsymbol{C}_T^{-1}(\boldsymbol{p}_t^d - \boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta}; \zeta_t))), \tag{11}$$

where $\boldsymbol{C}_T$ is some consistent estimate of $\boldsymbol{\Lambda}_0(\boldsymbol{\theta}_0)$ and $\boldsymbol{\theta}$ is a candidate parameter vector defined within the compact subspace $\boldsymbol{\Theta}$.

We come now to the central methodological contribution of the paper. Estimation based on the sample moments of equation (11) requires knowledge of equilibrium prices at the plant-area level (i.e., $\boldsymbol{H}(\boldsymbol{\theta}; \zeta_t)$). Yet the data generating process provides only prices that are aggregated and measured with error. The solution to this dilemma lies in numerical approximations to equilibrium. Conceptually, it is possible to *compute* the equilibrium price vector for any number of candidate parameter vectors, and then identify the candidate parameter vector that minimizes the "distance" between the aggregated equilibrium price vectors and the data. The power of modern computers makes this procedure feasible given a convenient distribution of the composite error term (ν_{ij}^* in equation (7)). In our application, we are typically able to numerically compute a vector, call it $\widetilde{\boldsymbol{H}}(\boldsymbol{\theta}; \zeta_t)$, that satisfies the first-order conditions of equation (6) to computer precision in a matter of seconds.

The GMM estimate that utilizes these numerical approximations is:

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \frac{1}{T}\sum_{t=1}^{T}[\boldsymbol{p}_t^d - \boldsymbol{S}(\widetilde{\boldsymbol{H}}(\boldsymbol{\theta};\zeta_t))]'\boldsymbol{C}_T^{-1}[\boldsymbol{p}_t^d - \boldsymbol{S}(\widetilde{\boldsymbol{H}}(\boldsymbol{\theta};\zeta_t))]. \tag{12}$$

We think it is intuitive to think of the estimation procedure as combining an outer loop and an inner loop. In the outer loop, the objective function is minimized over the parameter space, whereas in the inner loop equilibrium is computed numerically for each candidate parameter vector considered. This structure makes our estimator broadly analogous to other estimators developed for discrete static games (e.g., Bajari, Hong, and Ryan (2008)), non-strategic dynamic games (e.g., Rust (1987)) and certain strategic dynamic games (e.g., Goettler and Gordon (2009), and Gallant, Hong, and Khwaja (2010)), in the sense that each requires the repeated computation of equilibrium.

4.2 Asymptotic properties

The asymptotic properties of the GMM estimator are unclear without further assumptions, which we develop now:

Assumption A5: *A unique Bertrand-Nash equilibrium exists, and the prices that support it are strictly positive. Formally, for any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ there exists a vector $\boldsymbol{p}_1 \in \mathbb{R}_+^{JN}$ such that $\boldsymbol{f}(\boldsymbol{p_1};\zeta_t,\boldsymbol{\theta}) = \boldsymbol{0}$. Further, $\boldsymbol{f}(\boldsymbol{p_1};\zeta_t,\boldsymbol{\theta}) = \boldsymbol{f}(\boldsymbol{p_2};\zeta_t,\boldsymbol{\theta}) = \boldsymbol{0} \leftrightarrow \boldsymbol{p_1} = \boldsymbol{p_2}$.*

A5 ensures that the GMM objective function is well-behaved.[11] We suspect that uniqueness alone may suffice if, for instance, the econometrician can compute multiple equilibria and select the equilibrium closest to the data (e.g., as in Bisin, Moro, and Topa (2010)). We defer the evaluation of such possibilities to further research. The following lemma clarifies that, given the assumptions of the model, small changes to the parameter vector do not produce large jumps in the objective function:

Lemma 1: *The function $\boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta};\zeta_t))$ is continuously differentiable in $\boldsymbol{\theta}$ and $\boldsymbol{y}_t$ for $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, where $\boldsymbol{y}_t$ is the vector representation of ζ_t.*

Proof. See appendix A. □

[11] Recent theoretical contributions demonstrate that A5 holds for two special cases of our model: nested logit demand, convex marginal costs, and single-plant firms (Mizuno 2003), and logit demand, sufficiently increasing marginal costs, and multi-plant firms (Konovalov and Sándor 2010). The assumption is not satisfied generally (e.g., Caplin and Nalebuff (1991)).

Assumption A6: *The parameter vector* $\boldsymbol{\theta}_0$ *is globally identified in* $\boldsymbol{\Theta}$*. Formally,* $E[\boldsymbol{p}_t^d - \boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta};\zeta_t))|\zeta_t] = \boldsymbol{0} \leftrightarrow \boldsymbol{\theta} = \boldsymbol{\theta}_0$.

A6 could be violated even if parameters of the model would be globally identified given disaggregate data (i.e., even if $E[\boldsymbol{p}_t^\epsilon - \boldsymbol{H}(\boldsymbol{\theta};\zeta_t)|\zeta_t] = \boldsymbol{0} \leftrightarrow \boldsymbol{\theta} = \boldsymbol{\theta}_0$). Such a scenario may be more likely when aggregation is particularly coarse. Empirically, it may be possible to evaluate (imperfectly) the potential for this sort of aggregation problem using artificial data experiments, and we develop one such test in our application.

The asymptotic properties of the GMM estimator follow directly from A1-A6 and the other assumptions placed on the data generating process:

Theorem 1: *Under A1-A6 and certain regularity conditions enumerated in the appendix,*

i) $\widehat{\boldsymbol{\theta}} \rightarrow^p \boldsymbol{\theta}_0$ *and*

ii) $\sqrt{T}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \rightarrow^d N(\boldsymbol{0}, \boldsymbol{V})$,

where $\boldsymbol{V} = (\boldsymbol{G}_0'\boldsymbol{C}_0\boldsymbol{G}_0)^{-1}\boldsymbol{G}_0'\boldsymbol{C}_0\boldsymbol{\Lambda}_0\boldsymbol{C}_0\boldsymbol{G}_0(\boldsymbol{G}_0'\boldsymbol{C}_0\boldsymbol{G}_0)^{-1}$ *and* $\boldsymbol{G}_0 \equiv -E[\partial\boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta};\zeta_t))/\partial\boldsymbol{\theta}']$.

Proof. See appendix A. □

4.3 Incorporating non-price data

The estimation strategy can be extended to exploit variation in other endogenous data, such as observations on production or consumption, that are often available to the econometrician. We focus on production data for expositional brevity; the other extensions are analogous. We assume the data are generated by:

$$\boldsymbol{q}_t^\epsilon = \boldsymbol{q}(\boldsymbol{H}(\boldsymbol{\theta}_0;\zeta_t);\zeta_t,\boldsymbol{\theta}_0)) + \boldsymbol{\epsilon}_t^*, \tag{13}$$

where $\boldsymbol{q}_t^\epsilon$ is a vector of length JN, and $\boldsymbol{\epsilon}_t^*$ is a vector of unobserved sampling errors. We define a linear function $\boldsymbol{R}: \mathbb{R}^{JN} \rightarrow \mathbb{R}^{L^*}$ that maps the plant-area quantities to the aggregate production vector, which we denote as $\boldsymbol{q}_t^d$. We assume that the aggregate sampling error is mean zero conditional on the exogenous data:

$$E[\boldsymbol{q}_t^d - \boldsymbol{R}(\boldsymbol{q}(\boldsymbol{H}(\boldsymbol{\theta}_0;\zeta_t);\zeta_t,\boldsymbol{\theta}_0))|\,\zeta_t] = \boldsymbol{0}. \tag{14}$$

The GMM estimate that incorporates these data is:

$$\widehat{\boldsymbol{\theta}}^* = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} \boldsymbol{p}_t^d - \boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta};\zeta_t)) \\ \boldsymbol{q}_t^d - \boldsymbol{R}(\boldsymbol{q}(\boldsymbol{H}(\boldsymbol{\theta};\zeta_t);\zeta_t,\boldsymbol{\theta})) \end{bmatrix}' \boldsymbol{D}_T^{-1} \begin{bmatrix} \boldsymbol{p}_t^d - \boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta};\zeta_t)) \\ \boldsymbol{q}_t^d - \boldsymbol{R}(\boldsymbol{q}(\boldsymbol{H}(\boldsymbol{\theta};\zeta_t);\zeta_t,\boldsymbol{\theta})) \end{bmatrix}, \tag{15}$$

where $\boldsymbol{D}_T$ is some positive definite matrix, and the relevant contemporaneous variance matrix is $\boldsymbol{\Lambda}_0^*(\boldsymbol{\theta}_0) \equiv E[\boldsymbol{S}(\boldsymbol{\epsilon}_t)\ \boldsymbol{R}(\boldsymbol{\epsilon}_t^*)|\zeta_t]E[\boldsymbol{S}(\boldsymbol{\epsilon}_t)\ \boldsymbol{R}(\boldsymbol{\epsilon}_t^*)|\zeta_t]'$. Under A5 and an appropriately modified A6, Theorem 1 extends and the estimate is consistent and asymptotically normal.

5 The Portland Cement Industry

5.1 The product

Portland cement is a finely ground dust that forms concrete when mixed with water and coarse aggregates such as sand and stone. Concrete, in turn, is an essential input to many construction and transportation projects because its local availability and lower maintenance costs make it more economical than substitutes such as steel, asphalt, and lumber (Van Oss and Padovani (2002)). The producers of portland cement adhere to strict industry standards that govern the production process. Aside from geographic considerations, product differentiation in the industry is minimal.[12]

Producers negotiate private contracts with their customers, predominately ready-mix concrete firms and large construction firms. Most contracts specify a mill (or "free-on-board") price for portland cement at the location of production. Customers are responsible for door-to-door transportation, which is an important consideration because portland cement is inexpensive relative to its weight.[13] This fact is well understood in the academic literature. For example, Scherer et al (1975) estimates that transportation would account for roughly one-third of total customer expenditures on a hypothetical 350-mile route between Chicago and Cleveland, and a 1977 Census Bureau study reports that more than 80 percent is transported within 200 miles.[14] More recently, Salvo (2010) presents evidence consistent with the importance of transportation costs in the Brazilian portland cement industry.

[12]The standards are maintained by the the American Society for Testing and Materials Specification for Portland Cement, and exist to protect the quality and reliability of construction materials.

[13]The bulk of portland cement is moved by truck, though some is sent by train or barge to distribution terminals and only then trucked to customers.

[14]Scherer et al (1975) examined more than 100 commodities and determined that the transportation costs of portland cement were second only to those of industrial gases. Other commodities identified as having particularly high transportation costs include concrete, petroleum refining, alkalies/chlorine, and gypsum.

Some details of the production process motivate the marginal cost specification we introduce below. Cement plants are typically adjacent to a limestone quarry. The limestone is fed into coal-fired rotary kilns that reach peak temperatures of 1400-1450° Celsius. The output of the kilns – clinker – is cooled, mixed with a small amount of gypsum, and ground in electricity-powered mills to form portland cement. Kilns operate at peak capacity with the exception of an annual maintenance period. When demand is particularly strong, managers sometimes forego maintenance at the risk of breakdowns and kiln damage. Consistent with these stylized facts, a recent report prepared for the Environmental Protection Agency identifies five main variable input costs of production: raw materials, coal, electricity, labor, and kiln maintenance (EPA (2009)).

5.2 The geographic space

We focus on California, Arizona, and Nevada over the period 1983-2003. We refer to these three states as the "U.S. Southwest" for expositional convenience. Figure 2 maps the geographic configuration of the industry in the U.S. Southwest circa 2003. Most plants are located along an interstate highway, nearby one or more population centers. Some firms own multiple plants but ownership is not particularly concentrated – the capacity-based Herfindahl-Hirschman Index (HHI) of 1260 is well below the threshold level that defines highly concentrated markets in the 1992 Merger Guidelines. The figure also plots the four customs offices through which foreign imports enters the region – San Francisco, Los Angeles, San Diego, and Nogales. Most cement imported into the region is produced by large, efficient plants located in Southeast Asia – the cost of freighter transport is large but not prohibitive. Exports are negligible because domestic plants are not competitive in the international market.

We observe only two plant closures and only one plant entry over 1983-2003, consistent with the substantial sunk costs of kiln construction (e.g., Ryan (2009)). Thus, the geographic configuration of the industry circa 2003 is representative of the sample period.

In Figure 3, we plot total consumption and production over the sample period, together with two measures of foreign imports.[15] These data suggest that the U.S. Southwest forms a "geographic space" in the spirit of the theoretical model, after one accounts for foreign

[15] Both consumption and production are highly cyclical, consistent with the role of cement as an input to construction projects. However, consumption is more cyclical due to the costliness of capacity adjustments (e.g., as documented in Ryan (2009)) and the gap between consumption and production increases in overall activity. Thus, while imported cement generally represents a small fraction of total consumption, it plays an important role when demand outstrips domestic capacity.

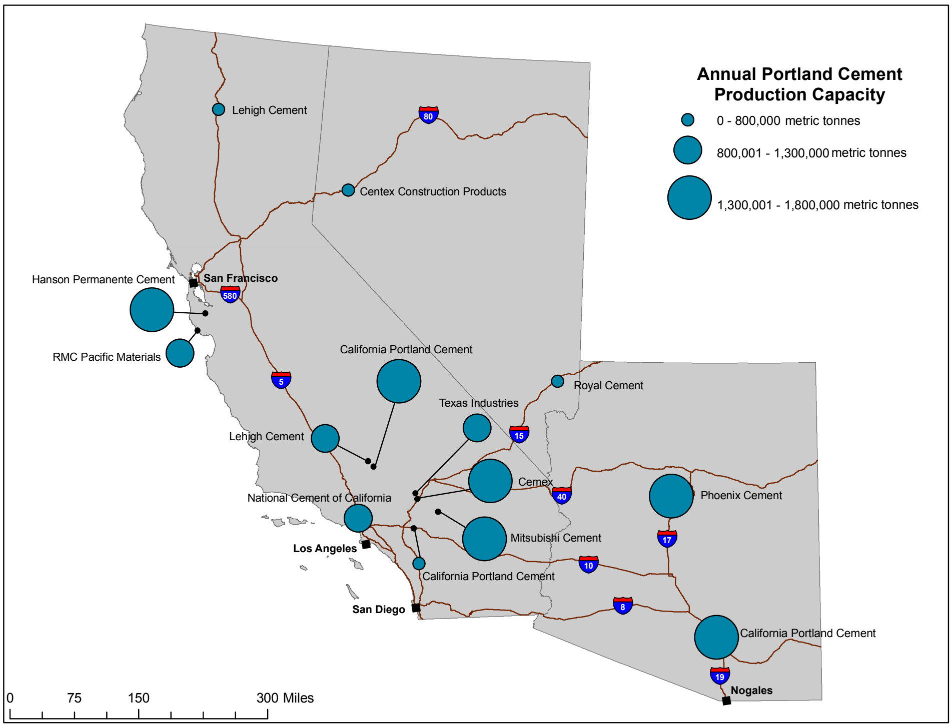

Figure 2: Portland Cement Production Capacity in the U.S. Southwest circa 2003.

imports. The similarity of the two imports measures we plot in Figure 3 – actual foreign imports and consumption minus production ("apparent imports") – reveals that net trade flows between the U.S. Southwest and other domestic regions are negligible. Other statistics published by the USGS are strongly suggestive that gross trade flows are also negligible. For instance, more than 98 percent of cement produced in southern California was shipped within the U.S. Southwest over 1990-1999, and more than 99 percent of cement produced in California was shipped within the region over 2000-2003. Outflows from Arizona and Nevada are unlikely because consumption routinely exceeds production in those states. And since net trade-flows between the U.S. Southwest and other domestic regions are insubstantial, these data points also imply that gross domestic inflows must also be insubstantial.

5.3 Data

We collect our endogenous data from the Minerals Yearbook, an annual publication of the U.S. Geological Survey (USGS). The Minerals Yearbook is based on an annual census of cement plants that collects detailed information on consumption, production, and mill prices.

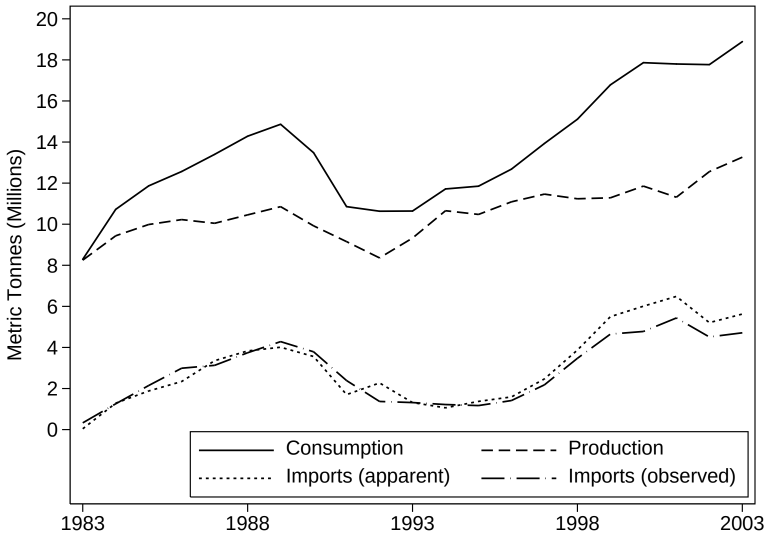

Figure 3: Consumption, Production, and Imports of Portland Cement. Apparent imports are defined as consumption minus production. Observed imports are total foreign imports shipped into San Francisco, Los Angeles, San Diego, and Nogales.

Census response rates are typically well over 90 percent, and the USGS staff imputes missing values for non-respondents based on historical and cross-sectional information.[16] The USGS aggregates the census data to the "regional" level before their publication in the Minerals Yearbook in order to protect the confidentiality of survey respondents. We observe the following endogenous data:

- Average mill prices (weighted by production) charged by plants in each of three regions: Northern California, Southern California, and a single Arizona-Nevada region.
- Total production by plants in the same three regions.
- Consumption in each of four regions: Northern California, Southern California, Arizona, and Nevada.

We also rely on the Minerals Yearbook for information on the price and quantity of portland cement that is imported into the U.S. Southwest.

We make use of more limited data on cross-region shipments from the California Letter, a second annual publication of the USGS. The level of aggregation varies over the

[16]The quality of the census has long generated interest among researchers. Other academic studies that feature USGS data include McBride (1983), Rosenbaum and Reading (1988), Jans and Rosenbaum (1997), Syverson and Hortaçsu (2007), and Ryan (2009).

sample period, some data are redacted to protect sensitive information, and no information is available before 1990. For instance, we observe shipments from producers in California (Northern and Southern) to consumers in Northern California over 1990-2003, but shipments from California to Nevada only over 2000-2003. There are 96 data points in total.

The Plant Information Survey (PIS), an annual publication of the Portland Cement Association, provides the geographic location of each portland cement plant as well as the annual kiln capacity and various other kiln characteristics. We collect county-level data from the Census Bureau on construction employment and residential construction permits, which help us specify the geographic distribution of potential demand. Finally, we collect data on diesel, coal, and electricity prices from the Energy Information Agency, data on average wages of durable good manufacturing employees from the BEA, and data on crushed stone prices from the USGS; we exploit state-level variation for all but the diesel data.

6 The Empirical Application

6.1 Specification

6.1.1 Supply

We let the marginal cost function of the domestic plants depend linearly on a set of exogenous cost-shifters and non-linearly on the level of capacity utilization at the plants:

$$c(Q_{jt}(\boldsymbol{\theta}_0); \zeta_t, \boldsymbol{\theta}_0) = \boldsymbol{w}'_{jt}\boldsymbol{\alpha} + \gamma \, 1\left\{\frac{Q_{jt}(\boldsymbol{\theta}_0)}{CAP_{jt}} > \nu\right\}\left(\frac{Q_{jt}(\boldsymbol{\theta}_0)}{CAP_{jt}} - \nu\right)^{\mu}, \tag{16}$$

where $\boldsymbol{w}_{jt}$ is a vector of input prices and CAP_j is total plant capacity. The constant portion of marginal costs can be derived from a Leontief production function (i.e., the factors of production are used in fixed proportions) and is consistent with the economics of portland cement production. The input prices we include are for coal, electricity, labor, and limestone. We let marginal costs increase in production once utilization exceeds some threshold ν. This treatment of capacity constraints, an innovation of Ryan (2009), imbeds the intuition that production near capacity creates shadow costs due to foregone kiln maintenance. The combination $\gamma(1-\nu)^{\mu}$ represents the marginal cost penalty associated with production at capacity. In practice, we find that it is difficult to estimate both γ and μ so we normalize μ to 1.5. (The marginal cost function is continuously differentiable for any $\mu > 1$.) The cost parameters to be estimated are $(\boldsymbol{\alpha}, \gamma, \nu) \in \boldsymbol{\theta}_0$.

We augment the theoretical model by letting domestic plants compete against a competitive fringe of foreign importers, which we denote as "plant" $J+1$. We place the fringe in geographic space at the four customs offices of the U.S. Southwest. Consumers pay the door-to-door cost of transportation from these customs offices. We rule out spatial price discrimination on the part of the fringe, consistent with perfect competition among importers, and assume that the import price is set exogenously (e.g., based on the marginal costs of the importers or other considerations). Thus, the supply specification is capable of generating the stylized fact that foreign importers provide substantial quantities of portland cement to the U.S. Southwest when demand is strong.

6.1.2 Demand

We express the indirect utility that consumers i receives from domestic plant j as follows:

$$u_{ijt} = \beta^c + \beta^p p_{jnt} + \beta^d MILES_{jn} * DIESEL_t + \nu^*_{ijt}, \tag{17}$$

where p_{jn} is the mill price per metric tonne, $MILES_{jn}$ is the miles (in thousands) between the plant and the centroid of the consumer's area and $DIESEL_t$ is a diesel price index that equals one in the year 2000. Hence, transportation costs increase linearly in distance and fuel costs, and the combination β^d/β^p is the cost per thousand tonne-miles given diesel prices at the 2000 level. We express the indirect utility received from the foreign importers as:

$$u_{i,J+1,t} = \beta^c + \beta^i + \beta^p p_{J+1,t} + \beta^d MILES_{J+1,n} * DIESEL_t + \nu^*_{i,J+1,t}, \tag{18}$$

where $MILES_{J+1,n}$ is the miles (in thousands) between the centroid of the consumer's area and the nearest customs office. The import-specific intercept is needed because the USGS data on import prices exclude duties. To be clear, the domestic prices are not observed in the data and must be computed as the solution to equation (6), given import prices that are exogenously-determined, non-discriminatory, and observed in the data. Finally, we normalize the mean value of the outside good to zero, so that $u_{i0t} = \nu^*_{i0t}$.

We assume that the distributions of the composite error terms (ν^*_{ijt}) generate a nested logit demand system in which the inside options (the domestic plants and the foreign imports) are in a different nest than the outside option. That is, we assume the composite error terms have i.i.d. extreme value distributions and define a parameter λ that characterizes the degree to which valuations of the inside options are correlated across consumers (e.g., as in Cardell (1997)). Valuations are perfectly correlated if $\lambda = 0$ and uncorrelated if $\lambda = 1$; the model collapses to a standard logit in the latter case. The demand parameters to be estimated are $(\beta^c, \beta^p, \beta^d, \beta^i, \lambda) \in \boldsymbol{\theta}_0$.[17]

The nested logit structure yields well-known analytical expressions for the quantity of cement that each plant sells to each area (i.e., $q_{jnt}(\boldsymbol{p_{nt}}; \zeta_t, \boldsymbol{\theta})$) and helps make estimation feasible from a computational standpoint. Nonetheless, the structure introduces some tension between the theoretical model and the empirical specification. Recall that the composite error term ν^*_{ijt} incorporates both an idiosyncratic preference shock and the consumer-specific deviation from mean distance (e.g., equation (7)). Since the deviation from mean distance is not independently distributed neither is the composite error.[18] The relevance of this discrepancy depends on how much of the variation in the composite error term is due to variation in deviations from mean distance. There should be less tension between the theoretical model and the empirical specification when areas are small, and more tension when areas are large or preference shocks are degenerate (e.g., as in the "pure characteristics model").

6.1.3 Areas and potential demand

We define 90 consumer areas based on the counties of the U.S. Southwest. The choice implies relatively fine spatial price discrimination and enables us to model the geographic distribution of demand using commonly-available data at the county level. We normalize potential demand using exogenous demand factors, following standard practice for discrete-choice systems (e.g., Berry, Levinsohn, and Pakes (1995), Nevo (2001)). The two factors we select are the number of construction employees and the number of new residential building permits. Thus, we implicitly assume that construction spending is unaffected by cement prices, consistent with the fact that that portland cement composes only a small fraction of total construction expenditures.[19]

[17] The substitution patterns between cement plants are characterized by the independence of irrelevant alternatives (IIA) within the inside good nest. IIA may be a reasonable approximation for our application. Portland cement is purchased nearly exclusively by ready-mix concrete plants and other construction companies. These firms employ similar production technologies and compete under comparable demand conditions. Thus, we are skeptical that meaningful heterogeneity exists in consumer preferences for plant observables (e.g., price and distance). Without such heterogeneity, the IIA property arises quite naturally – for example, the random coefficient logit demand model collapses to standard logit when the distribution of consumer preferences is degenerate.

[18] For a given consumer, the deviations from distance can be positively or negatively correlated. For instance, consider two plants located on either side of an area: a consumer that is closer to the first plant is farther from the second plant. But if the two plants are on the same side then a consumer that is closer to the first plant is also closer to the second plant.

[19] Syverson (2004) makes a similar argument for ready-mix concrete, which accounts for only two percent of total construction expenses according the 1987 Benchmark Input-Output Tables. The cost share of portland cement (an input to concrete) must be even lower.

To perform the normalization, we regress regional portland cement consumption on the demand predictors (aggregated to the regional level), impute predicted consumption at the county level based on the estimated relationships, and then scale predicted consumption by a constant of proportionality to obtain potential demand.[20] The results indicate that potential demand is concentrated in a small number of counties. In 2003, the largest 20 counties account for 90 percent of potential demand, the largest 10 counties account for 65 percent of potential demand, and the largest two counties – Maricopa County and Los Angeles County – together account for nearly 25 percent of potential demand. In the time-series, potential demand more than doubles over 1983-2003, due to greater activity in the construction sector and the onset of the housing bubble.

6.2 Estimation

We use a large-scale nonlinear equation solver developed in La Cruz, Martínez, and Raydan (2006) to compute equilibrium. The equation solver employs a quasi-Newton method and exploits simple derivative-free approximations to the Jacobian matrix; it converges more quickly than other algorithms and does not sacrifice precision. We define a numerical Bertrand-Nash equilibrium as a price vector for which $\frac{1}{J_t N} \parallel \boldsymbol{f}(\boldsymbol{p}\ ; \zeta_t, \boldsymbol{\theta}) \parallel < \delta$, where $\parallel \cdot \parallel$ denotes the Euclidean norm operator. The vector that defines numerical equilibrium given the 2003 data has $14 \times 90 = 1,260$ elements.[21]

We construct regional-level metrics based on the computed numerical equilibrium to compare the model predictions against the data. For notational convenience, we denote the elements of the equilibrium price vector as as $\widetilde{p}_{jnt}(\boldsymbol{\theta}; \zeta_t)$, and the corresponding quantities as $\widetilde{q}_{jnt}(\boldsymbol{\theta}; \zeta_t)$. We also define the sets $\aleph_r$ and $\jmath_r$ as the counties and plants, respectively, located

[20]The regression of regional portland cement consumption on the demand predictors yields an R^2 of 0.9786, which foreshadows an inelastic estimate of aggregate demand. Additional predictors, such as land area, population, and percent change in gross domestic product, contribute little additional explanatory power. We use a constant of proportionality of 1.4, which is sufficient to ensure that potential demand exceeds observed consumption in each region-year observation.

[21]We set δ =1e-13. Numerical error can propagate into the outer loop when the inner loop tolerance is substantially looser (e.g., 1e-7), which slows overall estimation time and can produce poor estimates. The inner loop tolerance is not unit free and must be evaluated relative to the price level.

in region r. Then the aggregated regional-level metrics take the form:

$$\begin{aligned}
\widetilde{p}_{rt}(\boldsymbol{\theta};\zeta_t) &= \sum_{j\in\mathcal{J}_r}\sum_{n}\frac{\widetilde{q}_{jnt}(\boldsymbol{\theta};\zeta_t)}{\sum_{j\in\mathcal{J}_r}\sum_n \widetilde{q}_{jnt}(\boldsymbol{\theta};\zeta_t)}\widetilde{p}_{jnt}(\boldsymbol{\theta};\zeta) \\
\widetilde{q}_{rt}(\boldsymbol{\theta};\zeta_t) &= \sum_{j\in\mathcal{J}_r}\widetilde{q}_{jnt}(\boldsymbol{\theta};\zeta_t), \\
\widetilde{c}_{rt}(\boldsymbol{\theta};\zeta_t) &= \sum_{j}\sum_{n\in\aleph_r}\widetilde{q}_{jnt}(\boldsymbol{\theta};\zeta_t),
\end{aligned} \tag{19}$$

where $\widetilde{p}_{rt}(\boldsymbol{\theta};\zeta_t)$ is the production-weighted average mill price, $\widetilde{q}_{rt}(\boldsymbol{\theta};\zeta_t)$ is total production, and $\widetilde{c}_{rt}(\boldsymbol{\theta};\zeta_t)$ is total consumption. We calculate regional prices and quantities for Northern California, Southern California, and the combined Arizona-Nevada region, and calculate regional consumption for Northern California, Southern California, Arizona, and Nevada.

We also exploit information on aggregated cross-region shipments to help identify the model.[22] We denote the quantity of shipments from plants in region r to consumers in region s as $\widetilde{q}^{s}_{rt}(\boldsymbol{\theta};\zeta_t)$. The shipments take the form:

$$\widetilde{q}^{s}_{rt}(\boldsymbol{\theta};\zeta_t) = \sum_{j\in\mathcal{J}_r}\sum_{n\in\aleph_s}\widetilde{q}_{jnt}(\boldsymbol{\theta};\zeta_t), \tag{20}$$

We calculate the quantity of portland cement produced by plants in California (both Northern and Southern) that is consumed in Northern California. The empirical analog is available over 1990-2003. We withhold other cross-region shipments from estimation because there are fewer data points and inclusion undermines the invertibility of the weighting matrix. Still, the withheld data provide natural checks on the model predictions.

The GMM estimate is:

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\sum_{t=1}^{T}\boldsymbol{e}_t(\boldsymbol{\theta};\zeta_t)'\boldsymbol{C}_T^{-1}\boldsymbol{e}_t(\boldsymbol{\theta};\zeta_t), \tag{21}$$

where $\boldsymbol{e}_t(\boldsymbol{\theta};\zeta_t)$ is a vector of empirical disturbances obtained by subtracting the aggregated model predictions specified by equations (20) and (20) from the corresponding data. We employ the usual two-step procedure to obtain consistent and efficient estimates (Hansen (1982)). We first minimize the objective function using $\boldsymbol{C}_T = \boldsymbol{I}$. We then estimate the contemporaneous variance matrix $\boldsymbol{\Lambda}_0$ and minimize the objective function a second time

[22]We find that the additional information is necessary to pin down the coal price coefficient in the marginal cost specificaion. This is unexpected but could be attributable to the high degree of correlation between coal prices and diesel prices over the sample period.

using the weighting matrix $\widehat{\boldsymbol{\Lambda}} \otimes \boldsymbol{I}$. We compute standard errors that are robust to both heteroscedasticity and arbitrary correlations among the error terms of each period, using the methods of Hansen (1982) and Newey and McFadden (1994).[23]

6.3 Identification

The use of aggregated data precludes point identification if multiple candidate parameters produce identical aggregate predictions despite having distinct disaggregate predictions. (This would violate A6.) To check for aggregation problems, we pair a vector of "true" parameters with 40 randomly-drawn sets of exogenous data. Both the parameters and the data are chosen to mimic our empirical application. For each of set of exogenous data, we compute equilibrium, generate the relevant aggregated data, and estimate using GMM. We argue that the model is reasonably identified if the GMM estimates are close to the true parameters.[24]

Table 1 shows the results of the artificial data experiment. Interpretation is complicated somewhat because we use non-linear transformations to constrain the some of coefficients (e.g., $\beta^p < 0$), and we defer details on these transformations to Appendix C. Nonetheless, it is clear that the means of the estimated coefficients are close to transformed true parameters. The means of the price and distance coefficients are within 6 percent and 11 percent of the truth, respectively. This precision is notable because the price and distance coefficients together determine the magnitude of transportation costs and thereby the degree of spatial differentiation (e.g., see the discussion of Equation 2). The other means of the estimated demand coefficients are somewhat farther from the truth. Among the marginal cost parameters, the mean estimated coefficients are accurate for the utilization threshold and the over-utilization cost but less accurate for the constant cost shifters. We conclude that the primary coefficients of interest (for spatial considerations) are likely well-identified but that

[23] Estimation of the contemporaneous variance matrix is complicated by the fact that we observe prices, production, and consumption over 1983-2003 but cross-region shipments over 1990-2003. We use methods developed in Srivastava and Zaatar (1973) and Hwang (1990) to account for the unequal numbers of observations.

[24] The exogenous data includes the plant capacities, the potential demand of counties, the diesel price, the import price, and two cost shifters. We randomly draw capacity and potential demand from the data (with replacement), and we draw the remaining data from normal distributions. Specifically, we use the following distributions: diesel price $\sim N(1, 0.28)$, import price $\sim N(50, 9)$, cost shifter 1 $\sim N(60, 15)$, and cost shifter 2 $\sim N(9, 2)$. We redraw data that are below zero and data that lead the estimator to nonsensical areas of parameter space. Throughout, we hold plant and county locations fixed to maintain tractability, and rely on the random draws of capacity, potential demand, and diesel prices to create variation in the distances between production capacity and consumers. Each artificial data set includes 21 draws on the exogenous data, with each draw representing a single time-series observation.

Table 1: Artificial Data Test for Identification

Variable	Parameter	Truth (θ)	Transformed ($\widetilde{\theta}$)	Mean Est	RMSE
Demand					
Cement Price	β^p	-0.07	-2.66	-2.51	0.66
Miles×Diesel Price	β^d	-25.00	3.22	2.86	0.59
Import Dummy	β^i	-4.00	-4.00	-6.07	1.23
Intercept	β^c	2.00	2.00	1.11	0.51
Inclusive Value	λ	0.09	-2.31	-1.73	0.54
Marginal Costs					
Cost Shifter 1	α_1	0.70	-0.36	-0.88	0.51
Cost Shifter 2	α_2	3.00	1.10	0.54	0.45
Utilization Threshold	ν	0.90	2.19	1.71	0.59
Over-Utilization Cost	γ	300.00	5.70	6.14	1.05

Results of GMM estimation on 40 data sets that are randomly drawn based on the "true" parameters listed. The parameters are transformed prior to estimation to place constraints on the parameter signs/magnitudes (see Appendix C). Mean Est and RMSE are the mean of the estimated (transformed) parameters and the root mean-squared error, respectively.

some skepticism of the other coefficients may be appropriate, especially with regard to the constant marginal cost shifters.

6.4 Multiple equilibria

We search for only a single equilibria in the inner loop, and problems may arise if multiple equilibria exist (this would violate A5). To provide some reassurance that the estimation procedure is reasonable, we conduct a Monte Carlo experiment and search for the existence of multiple equilibria. In particular, we compute equilibrium at eleven different starting points for thousands of randomly-drawn candidate parameter vectors. We then evaluate whether, for each given candidate parameter vector, the computed equilibrium prices are sensitive to the starting points.[25] More precisely, for each candidate parameter vector, we calculate

[25] We consider 300 parameter vectors for each of the 21 years in the sample, for a total of 6,300 candidate parameter vectors. For each $\theta_i \in \boldsymbol{\theta}$, we draw from the distribution $N(\widehat{\mu}_i, \widehat{\sigma}_i^2)$, where $\widehat{\mu}_i$ and $\widehat{\sigma}_i$ are the coefficient and standard error, respectively, reported in Table 2. We then compute the numerical equilibrium for each parameter vector, using eleven different starting vectors. We define the elements of the starting vectors to be $p_{jnt} = \phi \overline{p}_t$, where $\overline{p}_t$ is the average price of portland cement and $\phi = 0.5, 0.6, \ldots, 1.4, 1.5$. Thus, we start the equation solver at initial prices that are sometimes larger and sometimes smaller than the average prices observed in the data. The equal-solver computes numerical equilibria for 90 percent of the candidate vectors. See appendix C for a discussion of non-convergence in the inner-loop.

the standard deviation of each equilibrium price across the eleven starting points. (So there are 1,260 standard deviations for a typical equilibrium price vector of 1,260 plant-area elements.) The results indicate that the *maximum* standard deviation, over all candidate parameter vectors and all plant-area prices, is zero to computer precision. Thus, the Monte Carlo experiment finds no evidence of multiple equilibria. This may be unsurprising because, theoretically, uniqueness is ensured for two close cousins of our model: nested logit demand, convex marginal costs, and single-plant firms (Mizuno 2003), and logit demand, sufficiently increasing marginal costs, and multi-plant firms (Konovalov and Sándor 2010).

6.5 Key empirical relationships

Although the estimation routine relies on strong functional form assumptions on demand and marginal costs, it is nonetheless possible to visualize the key empirical relationships that drive the parameter estimates. We explore these relationships in Figure 4.

On the demand side, the price coefficient is primarily determined by the relationship between the consumption and price moments. In panel A, we plot cement prices and the ratio of consumption to potential demand ("market coverage") over the sample period. There is weak negative correlation, consistent with downward-sloping but inelastic aggregate demand. Next, the distance coefficient is primarily determined by (1) the cross-region shipments moment, and (2) the relationship between the consumption and production moments. We plot the gap between production and consumption ("excess production") for each region in panel B. In many years, excess production is positive in Southern California and negative elsewhere, consistent with inter-regional trade flows. The magnitude of these implied trade flows drives the distance coefficient. Interestingly, the implied trade flows are higher later in the sample, when the diesel fuel is less expensive.

On the supply side, the parameters on the marginal cost shifters are primarily determined by the price moments. In panel C, we plot the coal price, the electricity price, the durable-goods manufacturing wage, and the crushed stone price for California. Coal and electricity prices are highly correlated with the cement price (e.g., see panel A), consistent with a strong influence on marginal costs; inter-regional variation in input prices helps disentangle the two effects. It is less clear that wages and crushed stone prices are positively correlated with cement prices. Finally, the utilization parameters are primarily determined by (1) the relationship between the production moments (which determine utilization) and the consumption moments, and (2) the relationship between the production moments and the price moments. We explore the second source of identification in panel D, which shows

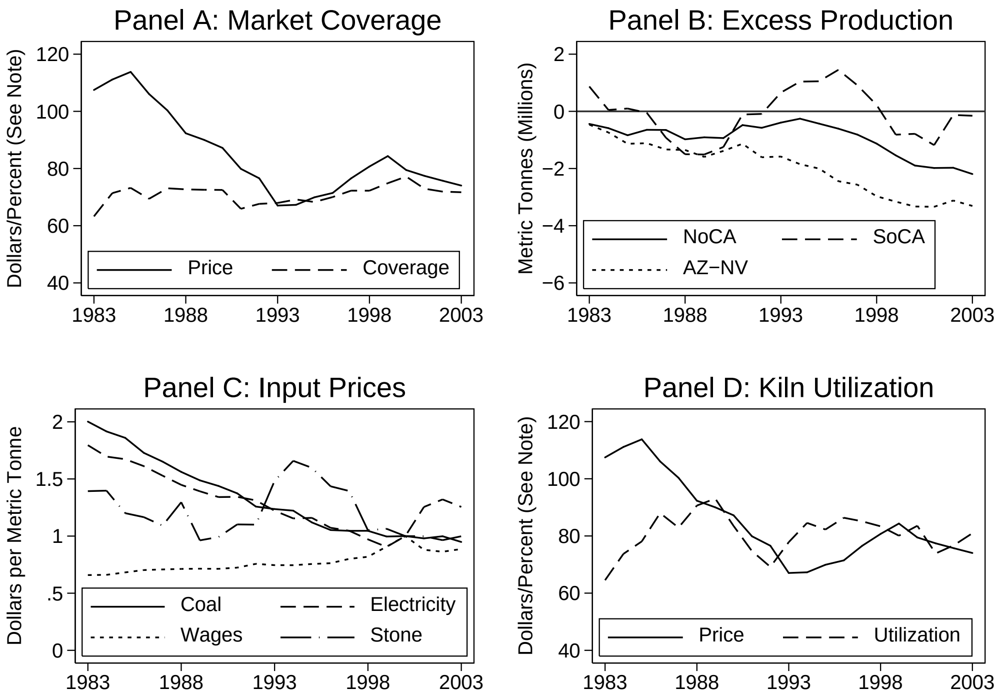

Figure 4: Empirical Relationships in the U.S. Southwest. Panel A plots average cement prices and market coverage. Prices are in dollars per metric tonne and market coverage is defined as the ratio of consumption to potential demand (times 100). Panel B plots excess production in each region, which we define as the gap between between production and consumption. Excess production is in millions of metric tonnes. Panel C plots average coal prices, electricity prices, durable-goods manufacturing wages, and crushed stone prices in California. For comparability, each time-series is converted to an index that equals one in 2000. Panel D plots the average cement price and industry-wide utilization (times 100).

cement prices and industry-wide utilization over the sample period. The two metrics are negatively correlated over 1983-1987 and positively correlated over 1988-2003.

7 Empirical Results

7.1 Demand estimates and transportation costs

Table 2 presents the parameter estimates of the GMM procedure. The price and distance coefficients are the two primary objects of interest on the demand side; both are negative and precisely estimated.[26] The aggregate elasticity implied by the price coefficient is -0.16 in the

[26]The other demand parameters take reasonable values and are precisely identified. The negative coefficient on the import dummy is likely due to the fact that observed import prices do not reflect the full price of imported cement (see Appendix D). The inclusive value coefficient suggests that consumer tastes for the

Table 2: Estimation Results

Variable	Parameter	Estimate	St. Error
Demand			
Cement Price	β^p	-0.087	0.002
Miles×Diesel Price	β^d	-26.42	1.78
Import Dummy	β^i	-3.80	0.06
Intercept	β^c	1.88	0.08
Inclusive Value	λ	0.10	0.004
Marginal Costs			
Coal Price	α_1	0.64	0.05
Electricity Price	α_2	2.28	0.47
Hourly Wages	α_3	0.01	0.04
Crushed Stone Price	α_4	0.29	0.31
Utilization Threshold	ν	0.86	0.01
Over-Utilization Cost	γ	233.91	38.16

GMM estimation results. Estimation exploits variation in regional consumption, production, and average prices over the period 1983-2003, as well as variation in shipments from California to Northern California over the period 1990-2003. The prices of cement, coal, and crushed stone are in dollars per metric tonne. Miles are in thousands. The diesel price is an index that equals one in 2000. The price of electricity is in cents per kilowatt-hour, and hourly wages are in dollars per hour. The marginal cost parameter ϕ is normalized to 1.5, which ensures the theoretical existence of equilibrium. Standard errors are robust to heteroscedasticity and contemporaneous correlations between moments.

median year, consistent with the conventional wisdom that materials such as steel, asphalt, and lumber are poor substitutes for portland cement in most construction projects. By contrast, the median firm-level elasticity of -5.70 is indicative of substantial price competition among the firms.

We estimate that consumers pay roughly \$0.30 per tonne mile, given diesel prices at the 2000 level.[27] Given the shipping distances that arise in numerical equilibrium, this translates into an average transportation cost of \$24.61 per metric tonne over the sample period – sufficient to account for 22 percent of total consumer expenditure. Transportation costs of this magnitude constrain the distance that cement can be shipped economically. We

different cement providers are highly correlated, inconsistent with the standard (non-nested) logit model.

[27]The ratio of the distance and price coefficients is the willingness-to-pay for proximity, incorporating transportation costs and any other distance-related costs (e.g., reduced reliability). We refer to the willingness-to-pay as the transportation cost for conciseness, although the two concepts may not be strictly equivalent. The specific calculation is $\frac{26.42}{0.087}\frac{index}{1000} = 0.3037$, where $index = 1$ in 2000.

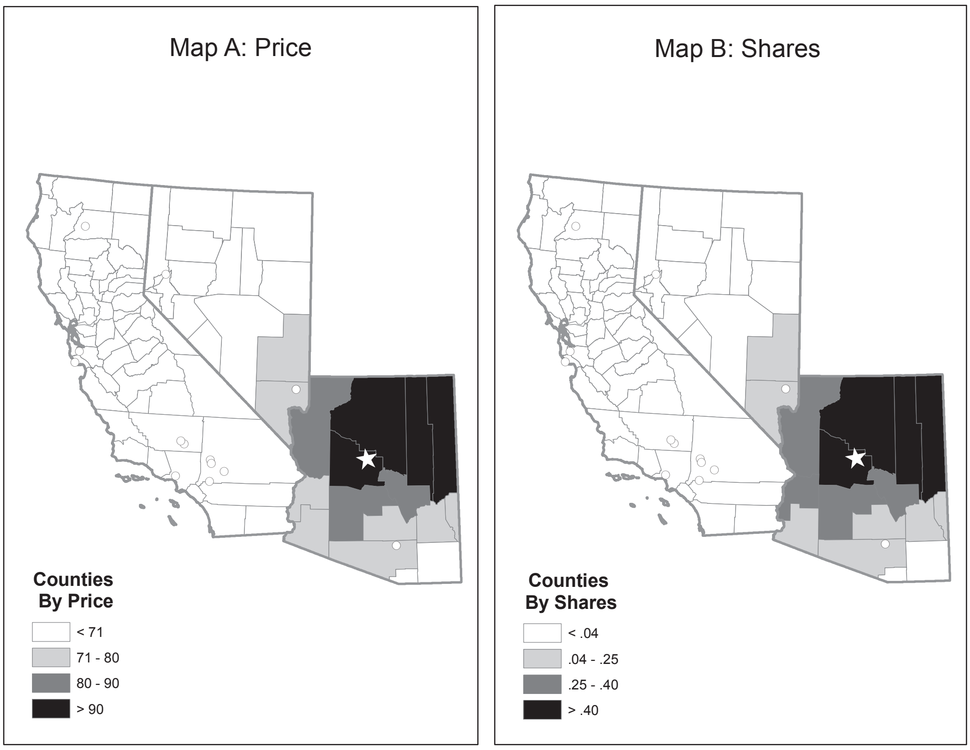

Figure 5: Equilibrium Prices and Market Shares for the Clarksdale Plant in 2003. The Clarksdale plant is marked with a star, and other plants are marked with circles.

find that portland cement is shipped an average of 92 miles, that 75 percent of portland cement is shipped under 110 miles, and that 90 percent is shipped under 175 miles.[28]

Firms appear to exercise some degree of localized market power. To illustrate, we map the prices and market shares of the Clarksdale plant that correspond to numerical equilibrium in Figure 5. We mark the location of the Clarksdale plant with a star, and mark other plants with circles. As shown, the Clarksdale plant captures more than 40 percent of the market in the central and northeastern counties of Arizona. It charges consumers in these counties its highest prices, typically $80 per metric tonne or more. Both market shares and prices are lower in more distant counties, and in many counties the plant captures less than one percent of demand despite steep discounts. The locations of competitors also influence market share and prices, though these effects are more difficult to discern.

We explore these relationships more rigorously with regression analysis. We regress prices and market shares on three independent variables: (1) the distance between the plant and the county, (2) the distance between the county and the nearest other domestic plant,

[28]The average shipping distance fluctuates between a minimum of 72 miles in 1983 and a maximum of 114 miles in 1998, and is highly correlated with the diesel price index.

and (3) the estimated marginal cost of the plant. Among plant-county pairs within 100 miles, a 10 percent reduction in distance is associated with prices and market shares that are 0.9 percent and 14 percent higher, respectively; and a 10 percent reduction in the distance separating the county from its the closest alternative is associated with prices and market shares that are 0.7 percent and 11 percent lower, respectively. Each of these patterns is statistically significant at the one percent level.[29]

7.2 Marginal cost estimates

We estimate marginal costs to be $69.40 in the mean plant-year (weighted by production). Of these marginal costs, $60.50 is attributable to costs related to coal, electricity, labor and raw materials, and the remaining $8.90 is attributable to high utilization rates. Integrating the marginal cost function over the levels of production that arise in numerical equilibrium yields an average variable cost of $51 million. Virtually all of these variable costs – 98.5 percent – are due to coal, electricity, labor and raw materials, rather than due to high utilization. Thus, although capacity constraints may have substantial affects on marginal costs, the results suggest that their cumulative contribution to variable costs can be minimal. Taking the accounting statistics further, we calculate that the average plant-year has variable revenues of $73 million and that the average gross margin (variable profits over variable revenues) is 0.32. As argued in Ryan (2009), margins of this magnitude may be needed to rationalize entry given the sunk costs associated with plant construction.[30,31]

Finally, we discuss the individual parameter estimates shown in Table 2, each of which deviates somewhat from production data available from the Minerals Yearbooks and EPA (2009). To start, the coal parameter implies that plants burn 0.64 tonnes of coal to produce one tonne of cement, whereas in fact plants burn roughly 0.09 tonnes of coal to produce each tonne of cement. The electricity parameter implies that plants use 228 kilowatt-hours per tonne of cement, whereas the true number is closer to 150. Each tonne of cement requires approximately 0.34 employee-hours yet the parameter on wages is essentially zero.

[29] We refer the readers to the working paper for more details on this regression.

[30] Lafarge North America, one of the largest domestic producers, reports an average gross margin of 0.33 over 2002-2004 in its public accounting records.

[31] Fixed costs are well understood to be important for production, as well. The trade journal *Rock Products* reports that high capacity portland cement plants incurred averaged $6.96 in maintenance costs per production tonne in 1993 (Rock-Products (1994)). Evaluated at the production levels that correspond to numerical equilibrium in 1993, this number implies that the average plant would have incurred $5.7 million in maintenance costs relative to variable profits of $17.7 million. Our results suggest that the bulk of these maintenance costs are best considered fixed rather than due to high utilization rates. Of course, the static nature of the model precludes more direct inferences about fixed costs.

Lastly, the crushed stone coefficient of 0.29 is too small, given that roughly 1.67 tonnes of raw materials are used per tonne of cement. We suspect that these discrepancies are due to measurement error in the data.[32] Alternatively, they may be due to a failure of identification (e.g., see Section 6.3) or due to the implicit assumption that plant productivity is fixed over the sample period – it seems clear that the renegotiation of onerous labor contracts improved productivity in the 1980s (e.g., Northrup (1989), Dunne, Klimek, and Schmitz (2009)).

7.3 Regression fits

One measure of an econometric model's viability is in its ability to fit the data.[33] In Figure 6, we plot observed consumption against predicted consumption (panel A), observed production against predicted production (panel B), and observed prices against predicted prices (panel C). Univariate regressions of the data on the predictions indicate that the model explains 93 percent of the variation in regional consumption, 94 percent of the variation in regional production, and 82 percent of the variation in regional prices. Thus, the model performs reasonably well in accounting for the variation in the endogenous data.

It is also telling to examine the model's out-of-sample predictions. In panel D, we plot observations on cross-region shipments against the corresponding model predictions. We use 14 of these observations in the estimation routine – the shipments from plants in California to consumers in northern California over 1990-2003 – but the remaining 82 data points are withheld from the estimation procedure and do not influence the estimated parameters. Even so, the model explains 98 percent of the variation in these data.

The quality of these fits is underscored by the rich time-series variation in the data due to macro-economic fluctuations. To illustrate, we aggregate the data and the model predictions across regions, and plot the resulting time-series in Figure 7. Panel A shows consumption, panel B shows production, panel C shows imports, and panel D shows average prices (imports are defined as production minus consumption). In each case, the model predictions mimic the inter-temporal patterns observed in the data. Univariate regressions of the data on the predictions explain 96 percent of the variation total consumption, 75 percent of the variation in total production, 76 percent of the variation in imports, and 91

[32] In particular, the coal prices in the data are free-on-board and do not reflect any transportation costs paid by cement plants; cement plants may negotiate individual contracts with electrical utilities that are not reflected in the data; the wages of cement workers need not track the average wages of durable-goods manufacturing employees; and cement plants typically use limestone from a quarry adjacent to the plant, so the crushed stone price may not proxy the cost of limestone acquisition (i.e., the quarry production costs).

[33] We are unaware of any statistical specification tests that are suitable for GMM with optimal instruments, which is exactly identified by construction.

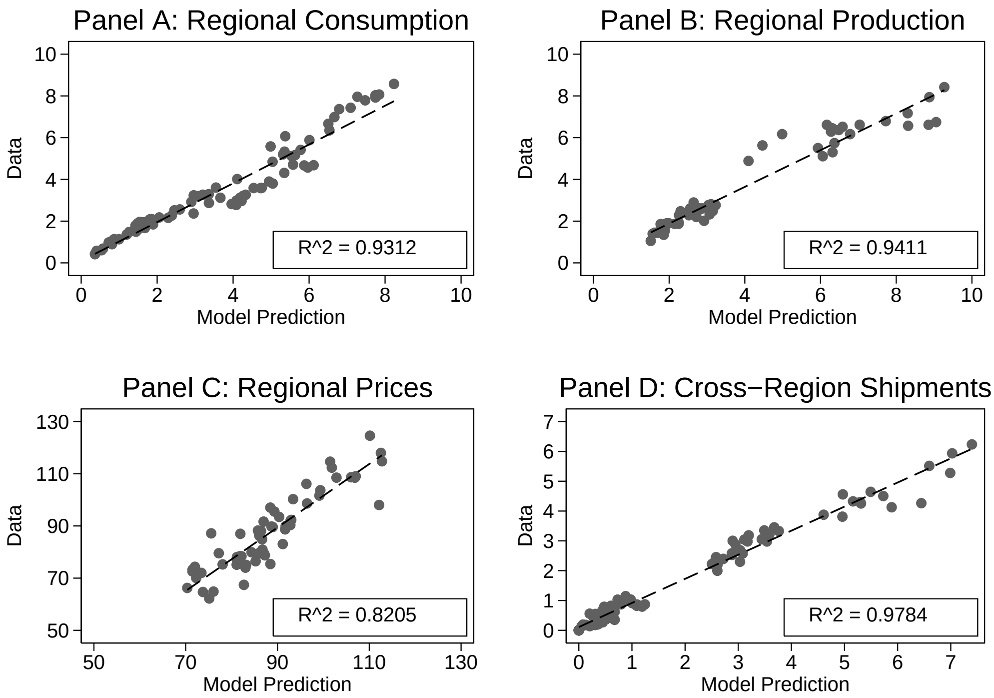

Figure 6: GMM Estimation Fits for Regional Metrics. Consumption, production, and cross-region shipments are in millions of metric tonnes. Prices are constructed as a weighted-average of plants in the region, and are reported as dollars per metric tonne. The lines of best fit and the reported R^2 values are based on univariate OLS regressions.

percent of the variation in average prices.[34]

7.4 An application to competition policy

The model and estimator may prove useful for a variety of policy endeavors. One potential application is merger simulation, an important tool for competition policy. We use counterfactual simulations to evaluate a hypothetical merger between Calmat and Gifford-Hill in 1986, when the firms together operated six plants and accounted for 43 percent of industry capacity in the U.S. Southwest.[35]

We map the distribution of consumer harm over the U.S. Southwest in Figure 8. In

[34] The model does not fully capture the fall in average prices over the 1980s and early 1990s. One plausible explanation is that the model does not account for the productivity improvements that occurred during the sample period (e.g., Northrup (1989), Dunne, Klimek, and Schmitz (2009)).

[35] We follow standard practice to perform the counterfactuals. We define an ownership matrix $\boldsymbol{\Omega^{post}(P)}$ that reflects the post-merger structure of the industry. We then compute the equilibrium post-merger price vector as the solution to Equation 6, substituting $\boldsymbol{\Omega^{post}(P)}$ for $\boldsymbol{\Omega(P)}$. Following McFadden (1981) and

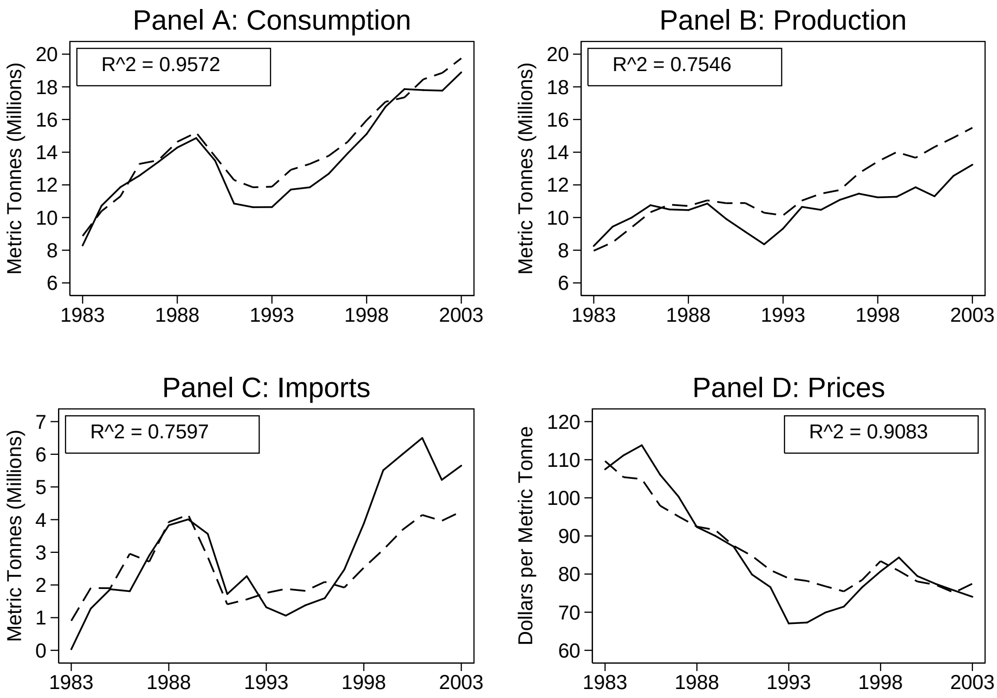

Figure 7: GMM Estimation Fits for Aggregate Metrics. The solid lines plot data and the dashed lines plot predictions. Consumption, production, and imports are in millions of metric tonnes. Imports are defined as production minus consumption. Prices are constructed as a weighted-average of the plant-county prices and are reported in dollars per metric tonne. The R^2 values are calculated from univariate regressions of the observed metric on the predicted metric.

panel A we focus on the effects of the merger, absent any divestitures. The total loss of consumer surplus is \$1.4 million which is small relative to pre-merger consumer surplus of \$239 million. Consumer harm is concentrated in the counties surrounding Los Angeles and Phoenix. Indeed, Maricopa County and Los Angeles County alone account for 60 percent of consumer harm and 10 counties account for more than 90 percent of the harm. We focus on potential remedies in panel B. We find that it is possible to eliminate 56 percent of the harm through the divestiture of a single plant. The divestiture of either the "Gifford-Hill 2" plant or the "Calmat 2" plant accomplishes this. As shown, however, these divestitures

Small and Rosen (1981), the change in consumer surplus due to the merger is:

$$\Delta CS = \sum_{n=1}^{N} \frac{\ln(1+\exp(\beta^c + \lambda I_{nt}^{pre})) - \ln(1+\exp(\beta^c + \lambda I_{nt}^{post}))}{\beta^p} M_n,$$

where I_n^{pre} is the inclusive value of the inside goods calculated using equilibrium pre-merger prices, I_n^{post} is the inclusive value calculated using equilibrium post-merger prices.

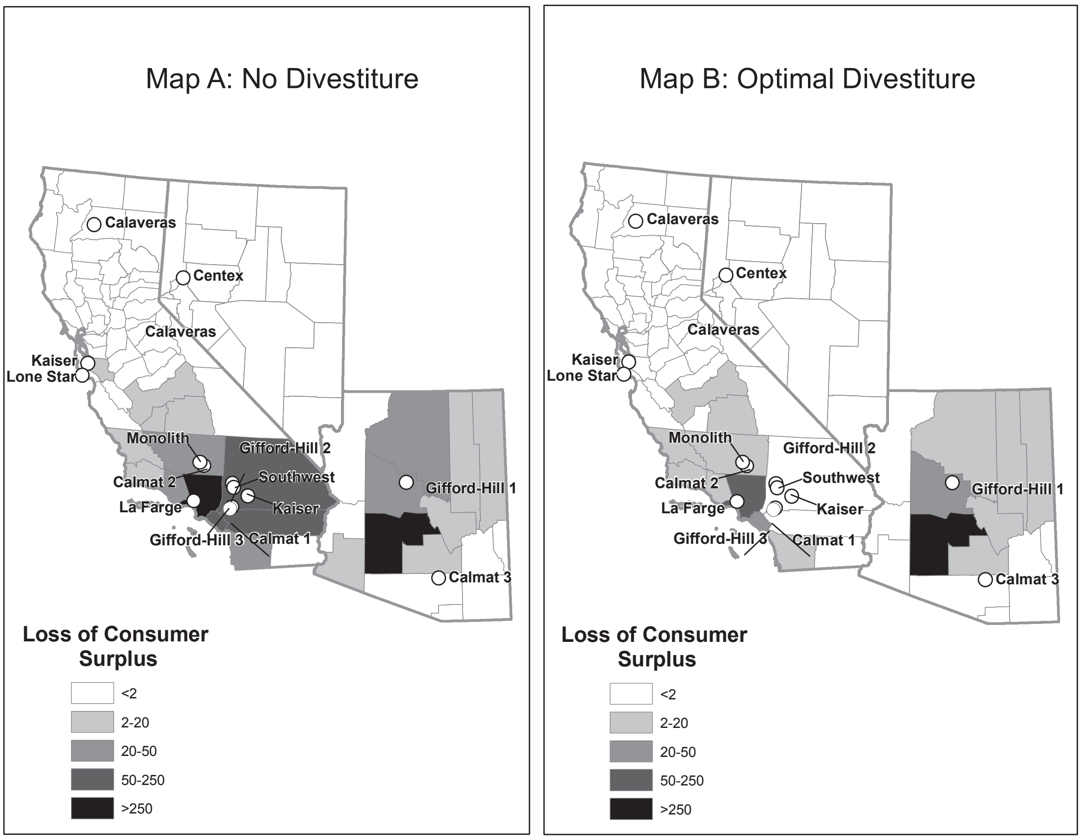

Figure 8: Loss of Consumer Surplus Due to a Hypothetical Merger between Calmat and Gifford-Hill

mitigates consumer harm in Southern California but do little to reduce harm in Maricopa County. Additional counterfactual exercises indicate that a two-plant divestiture is needed if this harm is to be mitigated as well.

7.5 Comparison to market delineation

In the introduction, we argue that the market delineation model imposes awkward theoretical assumptions. We now contrast some of our results to those of Ryan (2009), a recent paper that uses market delineation in a study of the portland cement industry. In particular, we point out that our approach generates distinctly different estimates of aggregate elasticity than does the market delineation approach. The discrepancy is consistent with the notion that our estimation strategy may sometimes provide more reasonable results than conventional approaches, and that these differences can be sizeable.[36]

[36]The discrepancy does not diminish the substantial contribution of Ryan (2009), which estimates an innovative dynamic discrete choice game and focuses primarily on the dynamic parameters; market delineation

Ryan makes the common assumptions that demand has constant elasticity and supply is Cournot within each market. He estimates the aggregate elasticity to be −2.96, which is quite different than our estimate of −0.16. The difference is entirely due to specification choices – the constant elasticity demand system produces an aggregate elasticity of −0.15 once housing permits are included as a control.[37] However, Ryan cannot use the inelastic estimate because, within the context of Cournot competition, it would imply that the firm elasticities are small to be consistent with profit maximization. This occurs because the Cournot model restricts each firm elasticity to be linearly related to the aggregate elasticity according the relationship $e_j = e/s_j$, where e_j, e, and s_j denote the firm elasticity, the aggregate elasticity, and the firm market shares, respectively. Further, Ryan cannot use the nested logit system to divorce the firm elasticities from the aggregate elasticity (as we do) because because logit models assume differentiated products whereas Cournot supply models assume homogenous products. Our takeaway is that our econometric strategy can lead to improved estimates by connecting the data to more realistic economic models.

8 Conclusion

The literature of the "new empirical industrial organization" focuses on the structural estimation of competition models and the recovery of the underlying parameters that guide firm and consumer decisions. Econometric innovations and greater computer power have improved our ability to link empirical correlations with sensible theoretical models of behavior. One area of particular interest has been the estimation of product differentiation models, as in Berry, Levinsohn, and Pakes (1995) and Nevo (2001). Yet geographic considerations – often critical drivers of differentiation – have received relatively little attention.

In providing an estimator for economic models of spatial price differentiation and spatial price discrimination, we hope to extend the reach of researchers to a number of questions that have long been emphasized in the theoretical literature. For instance, researchers could study the relationship between transportation costs and the intensity of competition, the welfare effects of spatial price discrimination, or the proper construction of antitrust markets. Though our empirical application is static, the estimator also could be used to define payoffs in strategic dynamic games. Such extensions could examine an array of interesting topics including entry deterrence, optimal location choice, and the effects of various government policies (e.g., carbon taxes or import duties) on welfare and the long-run location of

is used simply to determine the payoffs at different realizations of the state space.

[37] See Table 3 in Ryan (2009).

production. We are enthused by the breadth of opportunity.

References

Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press.

Bajari, P., H. Hong, and S. Ryan (2008). Identification and estimation of discrete games of complete information. Working paper, Duke University.

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

Bisin, A., A. Moro, and G. Topa (2010). The empirical content of models with multiple equilibria in economies with social interactions. Mimeo.

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

Collard-Wexler, A. (2009). Productivity dispersion and plant selection in the ready-mix concrete industry. *Mimeo*.

d'Aspremont, C., J. Gabszewicz, and J.-F. Thisse (1979). On Hotelling's "Stability in Competion". *Econometrica 47*, 1145–1150.

Davis, P. (2006). Spatial competition in retail markets: Movie theaters. *The RAND Journal of Economics 37*, 964–982.

Dunne, T., S. Klimek, and J. A. Schmitz (2009). Does foreign competition spur productivity? Evidence from post-WWII U.S. cement manufacturing. *Federal Reserve Bank of Minneapolis Staff Report*.

Economides, N. (1989). Stability in competition. *Journal of Economic Theory 47*, 178–194.

EPA (2009). *Regulatory Impact Analysis: National Emission Standards for Hazardous Air Pollutants from the Portland Cement Manufacturing Industry*. Prepared by RTI International.

Gallant, A. R., H. Hong, and A. Khwaja (2010). Dynamic entry with cross product spillovers: An application to the generic drug industry. Mimeo.

Goettler, R. and B. Gordon (2009). Does AMD spur Intel to innovate more? Mimeo.

Greenhut, J., M. Greenhut, and S.-Y. Li (1980). Spatial pricing patterns in the united states. *Quarterly Journal of Economics 94*, 329–350.

Hansen, L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*, 1029–1054.

Hotelling, H. (1929). Stability in competition. *Economic Journal 39*, 41–57.

Hwang, H.-s. (1990). Estimation of a linear SUR model with unequal numbers of observations. *Review of Economics and Statistics 72*.

Jans, I. and D. Rosenbaum (1997, May). Multimarket contact and pricing: Evidence from the U.S. cement industry. *International Journal of Industrial Organization 15*, 391–412.

La Cruz, W., J. Martínez, and M. Raydan (2006). Spectral residual method without gradient information for solving large-scale nonlinear systems of equations. *Mathematics of Computation 75*, 1429–1448.

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly Journal of Mathematics 2*, 164–168.

Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal of Applied Mathematics 11*, 431–331.

McBride, M. (1983, December). Spatial competition and vertical integration: Cement - concrete revisited. *American Economic Review 73*, 1011–1022.

McFadden, D. (1981). *Econometric Models of Probabilistic Choice*. MIT Press. Structural Analysis of Discrete Data, C.F. Manski and D. McFadden (eds.).

McManus, B. (Forthcoming 2009). Nonlinear pricing in an oligopoly market: The case of specialty coffee. *RAND Journal of Economics*.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*, 307–342.

Newey, W. K. and D. McFadden (1994). Large sample estimation and hypothesis testing. *Handbook of Econometrics 4*.

Northrup, H. R. (1989). From union hegemony to union disintegration: Collective bargaining in cement and related industries. *Journal of Labor Research 10*, 337–376.

Pesendorfer, M. (2003). Horizontal mergers in the paper industry. *RAND Journal of Economics 34*, 495–515.

Pinske, J., M. E. Slade, and C. Brett (2002). Spatial price competition: A semiparametric approach. *Econometrica 70*(3), 1111–1153.

Rock-Products (1994). Cement plant operating cost study. Overland Park, KS: Intertic Publishing.

Rosenbaum, D. and S. Reading (1988). Market structure and import share: A regional market analysis. *Southern Economic Journal 54*(3), 694–700.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurchner. *Econometrica 55*, 993–1033.

Ruud, P. A. (2000). *An Introduction to Classical Econometric Theory*. Oxford University Press.

Ryan, S. (2009). The costs of environmental regulation in a concentrated industry. Mimeo.

Salop, S. (1979). Monopolistic competition with outside goods. *Bell Journal of Economics 10*, 141–156.

Salvo, A. (2008). Inferring market power under the threat of entry: The case of the Brazilian cement industry. *Mimeo*.

Salvo, A. (2010). Trade flows in a spatial oligolopy: Gravity fits well, but what does it explain? *Canadian Journal of Economics 43*, 63–96.

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

Small, K. and H. Rosen (1981). Applied welfare economics with discrete choice methods. *Econometrica 49*, 105–130.

Srivastava, J. and M. K. Zaatar (1973). A Monte Carlo comparison of four estimators of the dispersion matrix of a bivariate normal population, using incomplete data. *Journal of the American Statistical Association 68*, 180–183.

Syverson, C. (2004, December). Market structure and productivity: A concrete example. *Journal of Political Economy 112*, 1181–1222.

Syverson, C. and A. Hortaçsu (2007). Cementing relationships: Vertical integration, foreclosure, productivity, and prices. *Journal of Political Economy 115*, 250–301.

Thisse, J.-F. and X. Vives (1988). On the strategic choice of spatial price policy. *American Economic Review 78*, 122–137.

Thomadsen, R. (2005). The effect of ownership structure on prices in geographically differentiated industries. *RAND Journal of Economics 36*, 908–929.

Van Oss, H. G. and A. Padovani (2002). Cement manufacture and the environment, Part I: Chemistry and technology. *Journal of Industrial Ecology 6*, 89–105.

Vogel, J. (2008, June). Spatial competition with heterogeneous firms. *Journal of Political Economy 116*(3), 423–466.

A Proofs

Proof of Lemma 1: We first show that $\boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta};\zeta_t))$ is continuously differentiable in $\boldsymbol{\theta}$ for $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. The proof is by contradiction. Suppose that $\boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta};\zeta_t))$ is not continuously differentiable at some parameter vector $\boldsymbol{\theta}_1 \in \boldsymbol{\Theta}$, i.e., that $\boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta};\zeta_t))/\partial\boldsymbol{\theta}'$ is discontinuous at $\boldsymbol{\theta}_1$. Then, by the linearity of $\boldsymbol{S}$ and the definition of discontinuity,

$$\lim_{\boldsymbol{\theta}^*\to\boldsymbol{\theta}_1^-} \left.\frac{\partial \boldsymbol{H}(\boldsymbol{\theta};\zeta_t)}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \neq \lim_{\boldsymbol{\theta}^*\to\boldsymbol{\theta}_1^+} \left.\frac{\partial \boldsymbol{H}(\boldsymbol{\theta};\zeta_t)}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}.$$

However, the function $\boldsymbol{f}(\boldsymbol{p},\boldsymbol{\theta};\zeta_t)$ is continuously differentiable in $\boldsymbol{p}$ and $\boldsymbol{\theta}$ by the assumptions placed on $\boldsymbol{q}(\boldsymbol{p},\boldsymbol{\theta};\zeta_t)$ and $\boldsymbol{c}(\boldsymbol{p},\boldsymbol{\theta};\zeta_t)$. It follows that $\partial\boldsymbol{f}(\boldsymbol{p},\boldsymbol{\theta};\zeta_t)/\partial\boldsymbol{\theta}'$ is continuous, i.e. that

$$\lim_{\boldsymbol{\theta}^*\to\boldsymbol{\theta}_1^-} \left.\frac{\partial \boldsymbol{f}(\boldsymbol{p},\boldsymbol{\theta};\zeta_t)}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \neq \lim_{\boldsymbol{\theta}^*\to\boldsymbol{\theta}_1^+} \left.\frac{\partial \boldsymbol{f}(\boldsymbol{p},\boldsymbol{\theta};\zeta_t)}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}.$$

Totally differentiating both sides, using the arbitrary price vector $\boldsymbol{H}(\boldsymbol{\theta};\zeta_t)$, yields

$$\begin{aligned}\lim_{\boldsymbol{\theta}^*\to\boldsymbol{\theta}_1^-} &\left(\frac{\partial \boldsymbol{f}(\boldsymbol{p},\boldsymbol{\theta};\zeta_t)}{\partial \boldsymbol{H}(\boldsymbol{\theta};\zeta_t)} \left.\frac{\partial \boldsymbol{H}(\boldsymbol{\theta};\zeta_t)}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} + \left.\frac{\partial \boldsymbol{f}(\boldsymbol{p},\boldsymbol{\theta};\zeta_t)}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}\right) \\ &= \lim_{\boldsymbol{\theta}^*\to\boldsymbol{\theta}_1^+} \left(\frac{\partial \boldsymbol{f}(\boldsymbol{p},\boldsymbol{\theta};\zeta_t)}{\partial \boldsymbol{H}(\boldsymbol{\theta};\zeta_t)} \left.\frac{\partial \boldsymbol{H}(\boldsymbol{\theta};\zeta_t)}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} + \left.\frac{\partial \boldsymbol{f}(\boldsymbol{p},\boldsymbol{\theta};\zeta_t)}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}\right).\end{aligned}$$

Since $\boldsymbol{f}(\boldsymbol{p},\boldsymbol{\theta};\zeta_t)$ is continuously differentiable in $\boldsymbol{p}$ and $\boldsymbol{\theta}$, it follows that

$$\lim_{\boldsymbol{\theta}^*\to\boldsymbol{\theta}_1^-} \left.\frac{\partial \boldsymbol{H}(\boldsymbol{\theta};\zeta_t)}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} = \lim_{\boldsymbol{\theta}^*\to\boldsymbol{\theta}_1^+} \left.\frac{\partial \boldsymbol{H}(\boldsymbol{\theta};\zeta_t)}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*},$$

which creates the contradiction. It remains to show that $\boldsymbol{S}(\boldsymbol{H}(\boldsymbol{\theta};\boldsymbol{y}_t))$ is continuously differentiable in $\boldsymbol{y}_t$ for $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, where $\boldsymbol{y}_t$ is the vector representation of ζ. The proof is obtainable by contradiction, using the same steps employed above, and we omit the explicit derivation for expositional brevity.

□

Proof of Theorem 1: We first place regularity conditions on the data generating process. Let $\boldsymbol{y}_t$ be the vector representation of the set ζ_t. We assume that $\{\boldsymbol{y}_t\}$ is a sequence of i.i.d. random vectors. We further assume that $\sup_{\theta\in\Theta}|\boldsymbol{p}_t^d - \boldsymbol{S}(\widetilde{\boldsymbol{H}}(\zeta_t,\boldsymbol{\theta}))| < \infty$, that $\sup_{\theta\in\Theta}|\partial\boldsymbol{H}(\zeta_t,\boldsymbol{\theta})/\partial\boldsymbol{\theta}| < \infty$, and that $\sup_{\theta\in\Theta}|[\boldsymbol{p}_t^d - \boldsymbol{S}(\widetilde{\boldsymbol{H}}(\zeta_t,\boldsymbol{\theta}))][\boldsymbol{p}_t^d - \boldsymbol{S}(\widetilde{\boldsymbol{H}}(\zeta_t,\boldsymbol{\theta}))]'| < \infty$. Amemiya (1985) proves that these conditions, along with the assumptions already introduced in the body of the text, imply the following properties:

(i) $\frac{1}{T}\sum_{t=1}^{T}[\boldsymbol{p}_t^d - \boldsymbol{S}(\widetilde{\boldsymbol{H}}(\boldsymbol{z}_t,\boldsymbol{\theta}))] \to^p E[\boldsymbol{p}_t^d - \boldsymbol{S}(\widetilde{\boldsymbol{H}}(\boldsymbol{z}_t,\boldsymbol{\theta}))]$ uniformly in $\boldsymbol{\theta} \in \boldsymbol{\Theta}$,

(ii) $\frac{1}{T}\sum_{t=1}^{T}[-\partial \boldsymbol{H}(\boldsymbol{z},\boldsymbol{\theta})/\partial\boldsymbol{\theta}] \to^p E[-\partial \boldsymbol{H}(\boldsymbol{z},\boldsymbol{\theta})/\partial\boldsymbol{\theta}]$ uniformly in $\boldsymbol{\theta} \in \boldsymbol{\Theta}$,

(iii) $\frac{1}{T}\sum_{t=1}^{T}[\boldsymbol{p}_t^d - \boldsymbol{S}(\widetilde{\boldsymbol{H}}(\boldsymbol{z}_t,\boldsymbol{\theta}))][\boldsymbol{p}_t^d - \boldsymbol{S}(\widetilde{\boldsymbol{H}}(\boldsymbol{z}_t,\boldsymbol{\theta}))]' \to^p \boldsymbol{\Lambda_0}(\boldsymbol{\theta})$ uniformly in $\boldsymbol{\theta} \in \boldsymbol{\Theta}$.

We refer readers to Theorem 4.2.1 (p. 116) in Amemiya (1985). We make the additional assumption that $\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{C}_T \to^p \boldsymbol{C}_0$, where $\boldsymbol{C}_0$ is some symmetric positive semidefinite matrix such that $E[\boldsymbol{p}_t^d - \boldsymbol{S}(\widetilde{\boldsymbol{H}}(\boldsymbol{z}_t,\boldsymbol{\theta}))] \notin \mathrm{Col}^{\perp}(\boldsymbol{C}_0)$ for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $\boldsymbol{\theta} \neq \boldsymbol{\theta_0}$. With properties (i)-(iii) and the convergence assumption on the weighting matrix, the proof is standard and we refer readers to Proposition 20 ("GMM Asymptotics," p. 546) in Ruud (2000).

□

B Summary Statistics

We provide selected summary statistics in Table 3. Some patterns stand out: First, substantial variation in each metric is available, both inter-temporally and across regions, to support estimation. Second, Southern California is larger than the other regions, whether measured by consumption or production. Third, consumption exceeds production in Northern California, Arizona, and Nevada; these shortfalls must be countered by cross-region shipments and/or imports. The observation that plants in these regions charge higher prices is consistent with transportation costs providing some degree of local market power. Finally, imports are less expensive than domestically produced portland cement. This discrepancy exists for two reasons: First, imports typically come in the form of clinker, which observes water from the air more slowly than cement. The clinker is ground into cement only after it clears customs. The import price does not include the grinding cost. Second, the import price does not include tariffs and duties, which are substantial. We include the import dummy in the demand specification to adjust for these factors.

C Estimation details

We minimize the objective function using the Levenberg-Marquardt algorithm (Levenberg (1944), Marquardt (1963)), which interpolates between the Gauss-Newton algorithm and the method of gradient descent. We find that the Levenberg-Marquardt algorithm outperforms

Table 3: Consumption, Production, and Prices

Description	Mean	Std	Min	Max
Consumption				
Northern California	3,513	718	2,366	4,706
Southern California	6,464	1,324	4,016	8,574
Arizona	2,353	650	1,492	3,608
Nevada	1,289	563	416	2,206
Production				
Northern California	2,548	230	1,927	2,894
Southern California	6,316	860	4,886	8,437
Arizona-Nevada	1,669	287	1050	2,337
Domestic Prices				
Northern California	85.81	11.71	67.43	108.68
Southern California	82.81	16.39	62.21	114.64
Arizona-Nevada	92.92	14.24	75.06	124.60
Import Prices [excludes duties and grinding costs]				
U.S. Southwest	50.78	9.30	39.39	79.32

Statistics are based on observations at the region-year level over the period 1983-2003. Production and consumption are in thousands of metric tonnes. Prices are per metric tonne, in real 2000 dollars. Import prices exclude duties. The region labeled "Arizona-Nevada" incorporates information from Nevada plants only over 1983-1991.

simplex methods such as simulated annealing and the Nelder-Mead algorithm, as well as quasi-Newton methods such as BFGS. We implement the minimization procedure using the nls.lm function in R, which is downloadable as part of the minpack.lm package.

We compute numerical equilibrium using Fortran code that builds on the source code of the dfsane function in R. The dfsane function implements the nonlinear equation solver developed in La Cruz, Martínez, and Raydan (2006) and is downloadable as part of the BB package. We find that Fortran reduces the computational time of the inner loop by a factor of 30 or more, relative to the dfsane function in R. The computation of equilibrium for each time period can be parallelized, which further speeds the inner loop calculations. The numerical computation of equilibrium takes between 2 and 12 seconds for most candidate parameter vectors when run on a 2.40GHz dual core processor with 4.00GB of RAM.

We use observed prices to form the basis of the initial vector in the inner loop computations, which limits the distance that the nonlinear equation solver must walk to compute numerical equilibrium. In practice, the equation solver occasionally fails to compute a numerical equilibrium at the specified tolerance level (1e-13) within the specified maximum number of iterations (600). The candidate parameter vectors that generate non-convergence in the inner loop tend to be less economically reasonable, and may be consistent with equilibria that are simply too distant from observed prices. When this occurs, we construct regional-level metrics based on the price vector that comes closest to satisfying our definition of numerical equilibrium.

To speed the inner loop computations, we re-express the first-order condition of 6 such that inversion of $\mathbf{\Omega}(\boldsymbol{p};\zeta,\boldsymbol{\theta})$ is avoided. The structure of the problem permits us to compute equilibrium separately for each period which is especially useful when multiple processors are available. We also note that when production is characterized by constant marginal costs one could further ease the computational burden of the inner loop by solving for equilibrium prices in each consumer area separately.

We constrain the signs and/or magnitudes of some parameters based on our understanding of economic theory and the economics of the portland cement industry, because some parameter vectors hinder the computation of numerical equilibrium in the inner loop. For instance, a positive price coefficient would preclude the existence of Bertrand-Nash equilibrium. We use the following constraints: the price and distance coefficients (β_1 and β_2) must be negative; the coefficients on the marginal cost shifters ($\boldsymbol{\alpha}$) and the over-utilization cost (γ) must be positive; and the coefficients on the inclusive value (λ) and the utilization threshold (ν) must be between zero and one. We use nonlinear transformations to implement the constraints. As examples, we estimate the price coefficient using $\widetilde{\beta}_1 = \log(-\beta_1)$ in the

GMM procedure, and we estimate the inclusive value coefficient using $\widetilde{\lambda} = \log\left(\frac{\lambda}{1-\lambda}\right)$. We calculate standard errors with the delta method.

D Data Details

We make various adjustments to the data in order to improve consistency over time and across different sources. We discuss some of these adjustments here, in an attempt to build transparency and aid replication. To start, we note that the California Letter is based on a monthly survey rather than on the annual USGS census, which creates minor discrepancies. We normalize the California Letter data prior to estimation so that total shipments equal total production in each year. The 96 cross-region data points include:

- CA to N. CA over 1990-2003
- CA to S. CA over 2000-2003
- CA to AZ over 1990-2003
- CA to NV over 2000-2003
- N. CA to N. CA over 1990-1999
- S. CA to N. CA over 1990-1999
- S. CA to S. CA over 1990-1999
- S. CA to AZ over 1990-1999
- S. CA to NV over 1990-1999
- N. CA to AZ over 1990-1999.

The (single) Arizona-Nevada region includes Nevada data only over 1983-1991. Starting in 1992, the USGS combined Nevada with Idaho, Montana and Utah to form a new reporting region. We tailor the estimator accordingly. Additionally, this region also includes information from a small plant located in New Mexico. We scale the USGS production data downward, proportional to plant capacity, to remove for the influence of this plant. Since the two plants in Arizona account for 89 percent of kiln capacity in Arizona and New Mexico in 2003, we scale production by 0.89. We do not adjust prices.

The portland cement plant in Riverside closed its kiln permanently in 1988 but continued operating its grinding mill with purchased clinker. We include the plant in the analysis over 1983-1987, and we adjust the USGS production data to remove the influence of the plant over 1988-2003 by scaling the data downward, proportional to plant grinding capacities. Since the Riverside plant accounts for 7 percent of grinding capacity in Southern California in 1988, so we scale the production data for that region by 0.93.

We exclude one plant in Riverside that produces white portland cement. White cement takes the color of dyes and is used for decorative structures. Production requires kiln temperatures that are roughly 50°C hotter than would be needed for the production of grey cement. The resulting cost differential makes white cement a poor substitute for grey cement.

The PCA reports that the California Cement Company idled one of two kilns at its Colton plant over 1992-1993 and three of four kilns at its Rillito plant over 1992-1995, and that the Calaveras Cement Company idled all kilns at the San Andreas plant following the plant's acquisition from Genstar Cement in 1986. We adjust plant capacity accordingly.

We multiply kiln capacity by 1.05 to approximate cement capacity, consistent with the industry practice of mixing clinker with a small amount of gypsum (typically 3 to 7 percent) in the grinding mills.

The data on coal and electricity prices from the Energy Information Agency are available at the state level starting in 1990. Only national-level data are available in earlier years. We impute state-level data over 1983-1989 by (1) calculating the average discrepancy between each state's price and the national price over 1990-2000, and (2) adjusting the national-level data upward or downward, in line with the relevant average discrepancy.