MARKET STRUCTURE AND PRODUCTIVITY:
A CONCRETE EXAMPLE

Chad Syverson

Working Paper **10501**

NBER WORKING PAPER SERIES

MARKET STRUCTURE AND PRODUCTIVITY: A CONCRETE EXAMPLE

Chad Syverson

Working Paper 10501
http://www.nber.org/papers/w10501

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2004

I thank seminar participants at the NBER Summer Institute, the Brookings Institution, the Center for Economic Studies, and numerous universities for their comments. I am grateful to Steve Levitt, Ariel Pakes, Mark Roberts, Balazs Szentes, and two anonymous referees for helpful suggestions and Kevin Clisham for answering many technical questions about the ready-mixed concrete industry. I am also indebted to John Haltiwanger, Rachel Kranton, Mike Pries, Plutarchos Sakellaris, and John Shea for their instruction and guidance. Funding from a Brookings Research Fellowship is appreciatively acknowledged. The research in this paper was conducted while the author was a research associate at the Center for Economic Studies, U.S. Bureau of the Census. Research results and conclusions expressed are those of the author and do not necessarily indicate concurrence by the Bureau of the Census or the Center for Economic Studies. This paper has been screened to ensure that no confidential information is released. The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

Market Structure and Productivity: A Concrete Example
Chad Syverson
NBER Working Paper No. 10501
May 2004
JEL No. L1, L6, D2

ABSTRACT

Many studies have documented large and persistent productivity differences across producers, even within narrowly defined industries. This paper both extends and departs from the past literature, which focused on technological explanations for these differences, by proposing that demand-side features also play a role in creating the observed productivity variation. The specific mechanism investigated here is the effect of spatial substitutability in the product market. When producers are densely clustered in a market, it is easier for consumers to switch between suppliers (making the market in a certain sense more competitive). Relatively inefficient producers find it more difficult to operate profitably as a result. Substitutability increases truncate the productivity distribution from below, resulting in higher minimum and average productivity levels as well as less productivity dispersion. The paper presents a model that makes this process explicit and empirically tests it using data from U.S. ready-mixed concrete plants, taking advantage of geographic variation in substitutability created by the industry's high transport costs. The results support the model's predictions and appear robust. Markets with high demand density for ready-mixed concrete – and thus high concrete plant densities – have higher lower-bound and average productivity levels and exhibit less productivity dispersion among their producers.

Chad Syverson
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
syverson@uchicago.edu

I. Introduction

Recent empirical work has left little doubt about the magnitude of plant-level productivity variation: it is enormous. This heterogeneity is also persistent. Perhaps surprisingly, much of the variation cannot be explained by differences between (even narrowly defined) industries. Studies reviewed in Bartelsman and Doms (2000), for example, have found 85^{th}-to-15^{th} total factor productivity percentile ratios between 2:1 and 4:1 within various four-digit SIC industries. A theoretical literature has arisen attempting to explain the sources of such diversity. The great majority of this research focuses on technological (i.e., supply-side) explanations, such as management influences, capital vintage effects, and R & D efforts.[1]

While supply-side effects are most certainly important, this paper focuses instead on the influence of the demand (i.e., output market) side. Specifically, it explores how product market characteristics can allow such large productivity differences (perhaps arising in part because of supply-side factors) to persist in long-run equilibrium. Key to the story is the ability of consumers to substitute the output of one supplier for another. The more difficult it is for consumers to switch between competing suppliers, the greater the productivity dispersion that can be sustained.

The role of the demand market, and substitutability specifically, in creating the large observed productivity differences becomes apparent when one considers how such wide efficiency variations can exist in equilibrium. If consumers were unencumbered by substitution barriers, production would be reallocated to a select few highly productive plants. Those plants who could produce output at lower cost than industry rivals would be able to grab additional market share by undercutting their opponents' prices without sacrificing profitability.

Such output and productivity patterns are not usually observed in the data, however. Virtually all industries, and indeed even markets within industries, exhibit widely varying producer productivity levels within them. Barriers to substitution across producers (i.e., various forms of product differentiation—be they spatial, physical, or brand-driven) can allow less productive plants to survive and even thrive in long-run equilibrium. Decreases in impediments to substitution, on the other hand, make it more difficult for low-productivity plants to profitably operate, truncating the equilibrium productivity distribution from below. The testable premise of

[1] Just a sampling includes Jovanovic (1982) and Ericson and Pakes (1995). See Bartelsman and Doms (2000) for a review of this literature.

this notion is that in markets where it is easy for industry consumers to switch suppliers, productivity distributions should exhibit higher minima, less dispersion, and higher central tendency than those in low-substitutability markets.

I test this mechanism within a single four-digit SIC industry, ready-mixed concrete (SIC 3273), and focus on a particular component of substitutability, spatial differentiation created by transport costs. The purpose of the paper, however, is not to give the final word on transport costs and productivity in a particular industry. Instead, I hope to show through a detailed case study how transport costs as well as other substitutability factors might impact productivity variation and levels throughout the economy.

The primary advantage of an industry case study is that it helps control for the influence of technology differences on productivity heterogeneity, isolating the demand-side impacts of interest. Additionally, focusing on an industry with a (relatively) physically homogeneous output that is subject to substantial transport costs clarifies the exercise of formally posing and testing the intuitive hypothesis above. The high transport costs imply the industry is actually a collection of quasi-independent geographic markets, all potentially subject to idiosyncratic demand movements. I take advantage of this across-market variation in the empirical tests. Ready-mixed concrete's homogeneity serves to isolate the source of product substitutability: since producers' outputs are physically comparable, transport-cost-driven spatial differentiation is what matters. The sharp focus on spatial substitutability is useful because it is arguably more easily measured than physical or brand-driven product differentiation. Spatial effects on productivity are also topics of specific interest to a considerable body of research.[2]

I model and empirically test a spatial competitive structure where increases in demand density (demand per unit area) in local markets truncate the producer productivity distribution from below. This implies dense markets will have higher minimum and average productivity levels and less productivity dispersion than low-density markets. Further, producers in higher-

[2] To see if the results of this case study hold more broadly, I investigate in Syverson (2004) how across-industry differences in measurable output substitutability factors are correlated with industries' plant-level productivity distribution moments. I find that manufacturing industries with lower transport costs, less physical product differentiation, and/or lower advertising intensities (all plausibly indicators of greater substitutability) do indeed tend to have less dispersed productivity distributions with higher averages than industries with more segmented output markets. In exchange for its broader focus, that study gives up some of the ability to control for the productivity effects of technological differences that the present empirical approach enjoys.

density markets will be larger on average and each serve a greater number of customers.[3]

Figures 1 and 2 offer preliminary evidence for this mechanism. They show kernel probability density estimates of the total factor productivity and size distributions for my sample of concrete plants. (The distributions are expressed in terms of deviations from the average across all plants in a given year.) Two distributions are plotted in each figure. These correspond to producers that are located in markets either above or below the median demand density level in my sample. (The former set of producers numbers roughly 8500, while the latter 4800. I precisely describe how plant productivity and size are measured and define markets and demand density below.) If the intuition above is correct, we should expect that the productivity distribution in high-density markets looks like a truncated version of the low-density distribution. This pattern is evident in Figure 1. At low productivity levels, the distribution of the high-density-market plants consistently has less weight than the low-density-market distribution. At high productivity levels, this pattern is reversed.[4] These visual patterns are confirmed by the computed means; the average of the high-density distribution is 0.051 log points greater (s.e. = 0.006) than that of the low-density distribution. While the comparative dispersion of the two distributions is more difficult to see in the figure, the prediction above holds in this regard as well. The standard deviation of the distribution of plants in high-density markets is 0.340, as compared to 0.375 for low-density producers (an F-test for equality of variance is easily rejected). Figure 2 shows that the predicted positive correlation between average plant size and density is also present in my sample, and indeed even more stark. The average-sized high-density-market plant is 0.454 log points (s.e. = 0.019) larger than the average producer in a low-density market. The implications of these two figures will be tested in more

[3] The mechanism through which these effects operate will be explained in detail below, but can be summarized as follows. A denser market requires more producers in a given area to serve it. Substitutability is greater in markets where producers are densely packed, because concrete buyers have access to more alternative producers. High substitutability corresponds with greater competitive pressures, forcing low-performing producers out of business. This truncates the long-run equilibrium plant-level productivity distribution from below and leads to the stated implications regarding the producer productivity and size distributions.

[4] Of course, the high-density distribution will not be an exact truncation of its low-density counterpart because of inherent randomness in the data (e.g., measurement error), and because I am pooling across markets with different truncation points. Interestingly, still, two features match attributes of a truncated distribution. As can be seen, the modes of the two TFP distributions are very close to each other. Furthermore, the quantiles of the high-density markets distribution are outside the 95 percent confidence interval of quantiles of the low-density distribution, *except* at very high TFP levels (above the 95th percentile). Left-truncated distributions of course become more similar closer to their right tails.

detail and with greater rigor below, but they are suggestive prima facie evidence in support of the link between substitutability and producer-level productivity and size distributions.

Besides extending the work explaining productivity (and size) differences within industries, this paper touches on other related topics. One is the rich and lengthy literature on the relationship between competition and productivity.[5] Conceptually, product substitutability and "competitiveness" are quite similar. Markets with greater substitutability are more competitive in the sense that their higher cross-price elasticities more greatly reward (punish) relatively low- (high-) cost producers in terms of market share. Hence the paper's implied positive link between substitutability and average productivity levels supports the literature's common (though not unanimous) notion that competition breeds efficiency.[6]

A point of contact with a separate literature arises from the model's combined implications for productivity and size. A curious between-producer form of scale economies is implied: producers in denser markets will be both larger and more efficient on average, even if there are *no* internal scale economies in production. The observed scale effect is instead the product of selective survivorship; less productive establishments are eliminated when markets become denser. Observably, this competition-driven selection process looks very much like the spatial agglomeration mechanisms discussed in the urban and trade literatures—producers in dense markets are more efficient. Interestingly, however, this process is distinct in that it does not rely on technological properties or externalities (such as internal increasing returns, Marshallian thick labor market effects, information spillovers, etc.) typically appealed to in the literature.

The paper is organized as follows. In the next section, I construct a theoretical framework that formalizes the intuitive premise above. The data used to test the theory's implications is then discussed in Section III. I test the model in Section IV and check the results

---

[5] This literature is much too large to cite comprehensively. Recent examples include Aghion and Howitt (1992), Nickell (1996), Melitz (2003), Raith (2003), and Schmitz (2004).

[6] One can remain agnostic about the specific source of productivity gains when competition is intensified. One branch of the competition-productivity literature focuses on "slack" or X-efficiency. That is, competition-spurred productivity growth occurs because producers are forced to take costly action to become more efficient, as in Raith (2003) for example. However, in the mechanism modeled here, productivity growth is instead achieved by selection across establishments with fixed productivity levels; less efficient producers are pushed out of the market. Both mechanisms are influenced by market competitiveness in theory, and both are likely to play a role in reality. Measuring the relative size of the contribution of each to determining productivity differences is beyond the scope of this paper, however.

for robustness in the following section. Section VI concludes.

II. Model

To formalize the story linking demand density, output substitutability, and the market productivity and size distributions, I construct a theoretical framework that incorporates consumers choosing among spatially differentiated products sold by heterogeneous-cost suppliers. The model offers testable implications regarding how the primary exogenous factor of interest, demand density, affects the (endogenously determined) equilibrium productivity and output distributions.

The framework extends the work of Salop (1979) to allow for heterogeneous producer costs. The model also adds asymmetric information among producers regarding their production costs, largely as an analytical convenience in constructing an equilibrium, but this is not necessary to obtain the tested empirical implications. Because I am concerned here with differences in productivity distributions across markets rather than intertemporal fluctuations within them, dynamics are not a primary concern. Thus I model a simple two-stage entry/production decision meant to capture long-run differences in outcomes across markets. While simple, the model shows in a straightforward manner how differences in spatial substitutability (arising from demand density variation) affect the shape of producer productivity and size distributions in markets.

A. Market Structure

A continuum of consumers is evenly distributed around a circle of unit circumference with a density of D consumers per unit length. Consumers have an inelastic demand for one indivisible unit of ready-mixed concrete, and will purchase if the price is less than their reservation value. The price faced by consumers is equal to the factory-door price set by the producer plus a transport cost that increases linearly in the distance from supplier to consumer. That is, $p' = p + tx$, where p' is the price paid by the consumer, x is the length of the arc between the plant and the customer, and t parameterizes transport costs. I assume for simplicity that reservation values are high enough to ensure that all consumers purchase in equilibrium. Thus the total quantity of concrete sold in the market is D. Demand density D is the exogenous variable of focus; I draw testable empirical implications from its effect on the equilibrium.

The supply side of the market is determined in a two-stage, simultaneous entry game. In the first stage, a large number of ex-ante identical potential entrants consider whether to attempt to gain entry access into the market. To do so, they must pay a sunk setup cost s, which is identical for all entrants. All producers choosing to pay s receive an idiosyncratic marginal cost draw c_i from a common distribution $g(c)$ with support $[0, c_u]$, where c_u is an arbitrary upper bound. The setup cost can be interpreted as resources spent drawing up a business plan, making initial inquiries into production possibilities, and other pre-production activities that would give a producer insight into its costs of production.[7] Those paying s learn their own cost draw but do not observe the cost draws of others.

In the second stage, those who have learned their costs decide whether to commence production, given the expected number and marginal costs of competitors.[8] Those choosing to produce pay a common fixed production cost f (which is also assumed to be irretrievable should production commence, say because of the irreversibility of investment, or because it captures the value of forgone alternative uses of the productive resources), are then placed randomly at evenly spaced locations on the circle (hence every location is identical in expectation), and set their factory-door price p. Consumers make their purchases given the resulting set of transport-inclusive prices.

[7] There is good reason to believe that producers sink resources into making entry decisions before learning their type. Substantial empirical evidence (e.g., Dunne, Roberts, and Samuelson 1989; Baily, Hulten, and Campbell 1992; and Foster, Haltiwanger, and Krizan 2002) indicates that young plants have higher failure (exit) rates than incumbents. This suggests that entering producers do not typically know very well their own position vis-à-vis their competitors with regard to profitability components such as cost types. (The present model abstracts from this early-production period and instead has high-cost firms dropping out of the market before commencing operations.)

[8] The assumption that producers decide whether to operate without knowing their competitors' cost draws allows demand density's impact on the equilibrium productivity distribution to be obtained analytically. Given the spatial setup of the model, competition is "local" when producers' costs are common knowledge; that is, neighbors have a greater impact on optimal strategies than do market producers further away. While one can solve for optimal strategies in the common-knowledge construct (I show how to do this in the earlier version of this paper; see Syverson 2001), this requires computational simulations to solve the model, and it is furthermore not obvious what the equilibrium concept—necessary to pin down the number of producers—should be in such simulations. An asymmetric information setup transforms the pricing game into something similar to those commonly found in heterogeneous-producer monopolistic competition models, where a producer plays against an "industry average" (in this case, the other entrants' expected price). Thus competition is no longer local (at least in the strategic sense—obviously, neighbors' prices in this model affect *realized* outcomes more so than do the prices of more distant competitors), and simulations are no longer necessary to determine the equilibrium distribution of producer cost/productivity levels. I would further argue that assuming producers have private information about their costs may be realistic; data-gathering empirical economists know well how famously possessive firms are about their cost data. Below I conduct computational robustness checks to see if the obtained results hold up qualitatively to deviations from the present assumptions about cost information.

While the realized market shares and profits of this entry game are stochastic and depend upon particular producer cost and location realizations, one can analytically determine the *distribution* of the equilibrium set of producers. I show below that the key testable implication of the model—the link between demand density and selection-driven truncation of the ex-ante productivity distribution—holds regardless of the particular set of cost and location realizations. Truncation occurs before sales and profits are realized because high-cost producers have lower *expected* profits. I also show, by computationally simulating a modified version of the model, that the key qualitative results of the benchmark model are robust to changes in the specified information and timing structure.

To solve for the model's equilibrium, consider first the decision of whether or not to commence operations for a would-be producer that has already learned its cost draw. The producer will operate in the market if it expects to earn positive profits from doing so. Since prospective entrants do not know others' cost draws, potential competitors are equivalent in expectation. They must therefore decide whether to operate based on the profits resulting from competing in a market against an expected number of n competitors charging the same expected price.

Because 1) all consumers purchase in equilibrium, and 2) an assumption on the parameters that will be discussed in detail below, there will be between each pair of producers a consumer indifferent to purchasing from either.[9] The particular location of this consumer depends, of course, on the prices of the two plants and transport costs. For any two neighboring plants i and j (which are a distance $1/n$ apart, where n is the equilibrium number of producers in the market), the indifferent consumer is located at a distance $x_{i,j}$ from i, where $x_{i,j}$ solves

$$p_i + tx_{i,j} = p_j + t\left(\frac{1}{n} - x_{i,j}\right), \quad (1)$$

and p_i and p_j are the factory-door prices set by producers i and j, respectively. This equation can be solved to recover $x_{i,j}$ explicitly. The total quantity sold by a producer between rivals j on one side and k on the other is then $(x_{i,j} + x_{i,k})D$, where $x_{i,k}$ is similarly defined.

[9] Restricting parameter values to ensure an indifferent consumer between each pair of producers greatly simplifies analysis of the equilibrium. It eliminates from producers' optimal pricing decisions the possibility that a realized price difference between neighboring producers is so large that some producers would capture all of their neighbors' customers—on both sides. (The demand discontinuity inherent to linear transport costs [see Salop 1979] ensures that no producer will ever have zero customers on one side and positive sales on the other.)

However, because producers do not know their rivals' costs or prices when making entry decisions, $E(x_{i,j}) = E(x_{i,k})$:

$$E(x_{i,j}) = E(x_{i,k}) = E(x_i) = \frac{E(p) - p_i}{2t} + \frac{1}{2}E\left(\frac{1}{n}\right). \quad (2)$$

$E(p)$ is the expected price charged by other entrants, and $E(1/n)$ is the expected reciprocal of the number of producers, taking equilibrium strategies as given. The expected profit of a producer setting its factory-door price equal to p_i is then

$$E(\pi_i) = 2E(x_i)(p_i - c_i)D - f = \left(\frac{E(p) - p_i}{t} + E\left(\frac{1}{n}\right)\right)(p_i - c_i)D - f. \quad (3)$$

Maximizing this expression with respect to p_i yields the producer's optimal price given its expectations about the equilibrium market outcome:

$$p_i = \frac{1}{2}c_i + \frac{1}{2}E(p) + \frac{t}{2}E\left(\frac{1}{n}\right). \quad (4)$$

Not surprisingly, the optimal price increases in the producer's own realized cost, its expectation of its competitors' prices, and the expectation of n^{-1}. This expression can also be used to compute rivals' expected prices in terms of the expected cost draw among equilibrium producers:

$$E(p) = E(c) + tE\left(\frac{1}{n}\right). \quad (5)$$

Substituting (4) and (5) back into (3) yields markups as well as maximized expected market share and profits in terms of parameters, the producer's own cost, the expected cost of its rivals, and the expected (reciprocal of the) number of producers in the market:

$$p_i - c_i = \frac{1}{2}E(c) + \frac{t}{2}E\left(\frac{1}{n}\right) - \frac{1}{2}c_i \quad (6a)$$

$$2E(x_i) = \frac{1}{2t}E(c) + E\left(\frac{1}{n}\right) - \frac{1}{2t}c_i \quad (6b)$$

$$E(\pi_i) = \frac{D}{4t}\left(E(c) + 2tE\left(\frac{1}{n}\right) - c_i\right)^2 - f \quad (6c).$$

Clearly, price-cost markups and expected market shares are declining in c_i. Thus assuming a large enough c_u, there will be cost draws that imply negative expected profits from operations. Therefore a critical cost draw c^* exists such that entrants drawing $c_i > c^*$ choose not to produce.[10] This cutoff cost draw can be solved for by setting (6c) equal to zero:

$$E(\pi_i) = 0 \Rightarrow c^* = E(c) + 2tE\left(\frac{1}{n}\right) - \sqrt{\frac{4tf}{D}}. \quad (7)$$

Note that $E(c)$, the average cost draw among equilibrium producers, is itself an increasing function of c^*, since the latter determines a truncation point of $g(c)$. That is,

$$E(c) \equiv \int_0^{c^*} c \frac{g(c)}{G(c^*)} dc. \quad (8)$$

Substituting (7) back into (6c) yields operating profits (conditional on $c_i \leq c^*$) as a function of parameters, the endogenously determined c^*, and the producer's own cost:

$$E(\pi_i \mid c_i \leq c^*) = \frac{D}{4t}\left(c^* - c_i + \sqrt{\frac{4tf}{D}}\right)^2 - f \quad (9)$$

Note that it is possible that some producers with $c_i \leq c^*$ deciding to commence production may make negative profits ex-post—specifically, those relatively high-cost (but still below c^*) producers who enter only to find themselves next to competitors with lower-than-expected costs. However, because f is sunk, they are willing to remain in the market as long as they are selling positive quantities. Positive sales can be assured by assumptions on the model's parameters that I discuss in the appendix. These same assumptions ensure that there will be an indifferent consumer between any producer pair in equilibrium. As mentioned previously, this greatly simplifies the analysis.

What, then, determines c^*? To see this, consider the potential entrant's choice of whether to pay the sunk cost s in order to receive a cost draw c_i. The expected value of entry, V^e, (common to all producers since they are ex-ante identical) is equal to expected operating profits *before knowing one's own cost draw* minus the sunk entry cost. I impose a free entry condition: the (expected) number of entrants adjusts to set V^e to zero. That is,

$$V^e = \int_0^{c^*} \left[\frac{D}{4t}\left(c^* - c + \sqrt{\frac{4tf}{D}}\right)^2 - f\right] g(c) dc - s = 0. \quad (10)$$

The equilibrium c^* is the value that solves this expression given parameters D, t, f, s, and the ex-

[10] While some costs $c_i > c^*$ imply positive expected profits due to the quadratic form of (6c), such cost levels would also imply (nonsensical) negative markups and expected quantities sold.

ante productivity distribution $g(c)$.[11]

B. The Comparative Statics of Shifts in Demand Density

Cutoff Cost Level c.* The comparative static of primary interest is the sign of dc^*/dD; in other words, how do differences in demand density across markets affect the shape of equilibrium productivity distributions? Using the implicit function theorem,

$$\frac{dc^*}{dD} = \frac{-\partial V^e / \partial D}{\partial V^e / \partial c^*}, \quad (11)$$

where the numerator,

$$\frac{\partial V^e}{\partial D} = \int_0^{c^*} \left[\frac{1}{4t}\left(c^* - c + \sqrt{\frac{4tf}{D}} \right)^2 - \frac{1}{4t}\left(c^* - c + \sqrt{\frac{4tf}{D}} \right)\sqrt{\frac{4tf}{D}} \right] g(c)dc\,, \quad (12a)$$

simplifies to

$$\frac{\partial V^e}{\partial D} = \int_0^{c^*} \left[\frac{1}{4t}(c^* - c)^2 + \frac{1}{4t}(c^* - c)\sqrt{\frac{4tf}{D}} \right] g(c)dc > 0\,. \quad (12b)$$

This expression is positive because $c^* \geq c$ throughout the region of integration.

The demoninator of the implicit function theorem is given by

$$\frac{\partial V^e}{\partial c^*} = \left[\frac{D}{4t}\left(c^* - c^* + \sqrt{\frac{4tf}{D}} \right)^2 - f \right] g(c^*) + \frac{D}{2t}\int_0^{c^*} \left(c^* - c + \sqrt{\frac{4tf}{D}} \right) g(c)dc\,. \quad (13a)$$

Simplifying gives

$$\frac{\partial V^e}{\partial c^*} = \int_0^{c^*} \frac{D}{2t}\left(c^* - c + \sqrt{\frac{4tf}{D}} \right) g(c)dc > 0\,, \quad (13b)$$

because the first term in (13a) equals zero; allowing in a formerly marginally unprofitable producer by slightly increasing c^* has no impact on the value of entry.

Therefore $dc^*/dD < 0$; the upper bound of the producers' cost distribution decreases in

[11] The equilibrium c^* is unique. The logic is as follows. From (6c),expected profits conditional upon c_i are decreasing in c_i and strictly increasing in $E(1/n)$. For any given $E(1/n)$, the expected value of entry V^e for an entry policy x (that is, where entrants commit to commence production upon receiving any marginal cost draw $c_i \leq x$) is maximized at $x = c^*$, where c^* is defined as in (7). Since V^e is monotonic in $E(1/n)$, the unique equilibrium c^* is simply the value of this maximizing entry policy when $E(1/n)$ is such that $V^e = 0$. Interested readers can find formal existence and uniqueness proofs of similarly-styled equilibria—albeit with different assumptions regarding specifics of the demand and supply structures—in Asplund and Nocke (2003) and Melitz (2003).

demand density. High-cost (low-productivity) producers aren't profitable in dense markets. This is in accordance with the intuitive story forwarded in the introduction.

Two impacts of an increase in demand density (which is exogenous in the model, and as I will argue, empirically as well) are at work. The direct effect is through the increased sales and profits for any fixed number of market producers. This increase in potential profits in isolation would raise the expected value of entry. To preserve the equilibrium condition that V^e equals zero, $E(1/n)$ must fall in order to lower expected profits conditional upon c_i; see (6c). The additional entry decreases the average distance between producers, increasing substitutability and making it harder for high-cost producers to hold customers. Buyers formerly stuck with an inefficient supplier may now have a lower-cost option. In equilibrium, this raises the bar for successful entry into the market, yielding the cost/productivity truncation result.[12]

This progressive truncation of the cost distribution from above, coupled with the fact that measured productivity levels should be inversely related to costs, imply that average and minimum productivity levels (average and maximum costs) of producers in denser markets should be higher (lower) than in less dense markets. Also, assuming some additional regularities in the distribution, productivity dispersion should be lower in denser markets. These are the primary empirical implications I test below.

Firm Size. Increases in demand density induce entry, but do the number of producers grow at a proporational rate to the growth in market demand? In other words, what happens to average firm size as density rises? While the fact that the number of producers enters the model through the expectation of its reciprocal (and expectations are not commutative) precludes an analytical derivation of the elasticity of $E(n)$ with respect to D, the model still offers some guidance along these lines. As is evident in (6a), markups increase with $E(1/n)$. Thus the average markup falls in D because of the entry a rise in D induces. But since the expected value of entry in equilibrium is always zero, it must be that producers in dense markets—who have to cover the same sunk entry costs as producers in low-density markets—make up for the lower markup by selling more concrete on average. Therefore, we should expect average firm size to rise with

[12] One can show that, not surprisingly, a decrease in transport costs t implies a shift in c^* in the same direction as an increase in D. Both serve through different mechanisms to increase substitutability.

demand density.[13] I test this implication in the empirical work below.

C. Robustness of Comparative Statics

As stated above, the advantage of the current setup is that it allows analytical derivation of the empirically tested comparative statics. However, to do so it assumes away repricing and further entry after producers' sales are realized. It is conceivable that in reality, producers' prices would become common knowledge at this point, allowing producers' actual costs to be inferred. This would identify market locations "ripe" for new entry (i.e., those among relatively high-cost producers). There may be other potential producers hoping to take advantage of this identified weakness. And since these new entrants could identify the best location before entering, they could possibly be *less* efficient than the marginal entrant above. The current model rules out this instance because the final decision is made before prices (and therefore producer costs which can be inferred from them) are revealed.

To see how allowing ex-post entry affects the comparative statics derived above, I conduct a simple exercise that considers a stark case where, once all prices are realized, a new potential entrant can observe these and then choose the best location in which to enter the market.[14] In Syverson (2001) I show that when all producers' costs are known, operating profits for a producer with cost draw c_i equal

$$\pi_i = \frac{D}{t}\left[\frac{t}{n} + \sum_{j \neq i} s_{i,j} c_j - (1 - s_{i,i}) c_i\right]^2 - f, \quad (14)$$

where $s_{i,j}$ is the (i,j)-th element of a weighting matrix $\boldsymbol{S}$. This matrix, whose computation and properties are discussed in detail in Syverson (2001), captures the influence of competitors' costs on a producer's optimal price as well as the resulting market shares and profits.

I can use this expression to compute the highest cost draw c_i that an "ex-post" entrant choosing to enter at a specific location i could have and still earn nonnegative operating profits. Notice from the second term in the brakets in (14) that a producer's profits depend on a weighted

[13] The result that the number of producers grows less than proportionately to the size of the market has been found empirically in other industries by Bresnahan and Reiss (1991) and Campbell and Hopenhayn (2002).

[14] Here I assume that incumbents can reset their prices optimally against the entrant and each other. I also preserve the assumption of equal spacing of producers after the entrant begins operations for tractability. More precisely speaking, then, the new entrant is picking its neighbors rather than a specific location on the circle.

sum of its competitors' cost levels. Therefore the best location to enter for any new producer (regardless of its own cost draw) is that location i where this term is highest—i.e., where its closest competitors have the highest costs in the market. Setting (14) equal to zero and solving for c_i at this weakest-competitor location yields the highest cost level an ex-post entrant could have and still earn nonnegative operating profits. In some sense, this value is a conservative estimate—an upper bound—of the marginal costs that would be observed empirically, because it allows the ex-post entrant to locate next to the weakest competitors after identifying their location. It is additionally conservative because this analysis ignores whether such ex-post entry would even be profitable in expectation once the sunk entry cost s is figured in.

I compute two versions of this upper-bound cost by simulating the model repeatedly. Details of the simulation process are in the appendix. For a fixed demand density level and parameter set, 1000 market equilibria are determined as outlined above; potential producers draw costs from a common distribution, decide whether to produce, and (if so) set their prices according to the optimal behavior exposited in the model. Then the ex-post upper bound c_i is computed for each equilibrium using (14). The maximum and average of these upper bounds across the 1000 trials are then computed. As a check on the intuition given above regarding the relative growth of n with respect to D, I also compute the average producer size in the simulated outcomes. This simulation process is repeated over a range of demand density values.

The results of this exercise are shown in Figure 3. It plots for a range of demand density levels the maximum and average upper-bound marginal costs, the average producer size, and for comparison the c^* implied by the model. As can be seen, the negative relationship between demand density and the highest cost in the market remains. Furthermore, the simulations imply that, as supposed above, higher demand density leads to larger producers on average.

While simple, these simulation exercises do suggest that the basic insights of the model should remain in the presence of any (virtually inevitable) deviations between the mechanisms in the model and those existing in reality.

III. Data and Measurement Issues

A. Producer Productivity and Size

Computing producer-level productivity and size distribution moments are key to testing the model. I do so using data from U.S. ready-mixed concrete (SIC 3273) plants in the 1982, 1987, and 1992 Census of Manufactures (CM).[15] The CM microdata contain a wealth of information on plants' production activities. Importantly here, they also contain the state and county in which establishments are located, allowing me to place producers into defined geographic markets.

Each year of the quinquennial CM contains information on each of the roughly 5200 ready-mixed plants operating in the United States. However, very small plants (typically with fewer than five employees)—called Administrative Record (AR) establishments—have imputed data for most production variables. These AR plants amount to roughly one third all establishments, though their small size implies they compose a much smaller share of employment and output. Because of their imputed data, I exclude these plants from the main sample. However, I do observe their location, so I can count the total number of market producers for use when needed.

I augment the standard CM data with the accompanying CM Product data files. These auxiliary files contain information on the specific products (defined at the seven-digit SIC level) made by each establishment. This information includes the total value of shipments by product for each plant, and when measurable, product output in physical units. This physical output measure is available for almost all non-AR ready-mixed concrete plants (indeed, the industry is unusual in that it only contains a single seven-digit product). Thus I can measure producers' ready-mixed shipments either in dollars or cubic yards.[16]

This product data is valuable from a measurement standpoint. It allows me to use a physical output measure rather than relying on deflated revenue as most plant-level studies must do because of the dearth of plant-level price indices. Deflated-revenue output measures pose problems when prices vary across plants for reasons other than output quality (differences in

[15] In most of the empirical work, I do not distinguish between plants and firms. Most ready-mixed concrete plants in the U.S. during my sample period were single-unit firms (although this fraction has been falling over time). For example, 3749 firms controlled the 5319 ready-mixed plants operating in 1987. In earlier versions, I tested the empirical results for sensitivity to the prevalence of multi-plant firms (along with other technological factors that may vary across markets), and there were not large differences—see Syverson (2001). Moreover, productivity still varies across plants in multi-unit firms, and the selection mechanism exposited here can also be applied to determine which plants *within* a firm survive in equilibrium.

[16] In addition to AR plants, I also exclude those few plants who earn less than half of their revenue from ready-mixed concrete. (While plants' industry classifications are typically based on the product that comprises the their largest share of revenue, some that were classified as ready-mixed did not earn the majority of their revenue through ready-mixed sales.) However, most ready-mixed producers are quite specialized; the average revenue share of ready-mixed across all plants in my sample is 95.5 percent.

demand conditions across plants, say), because productivity measures then embody within-industry price variation. That is, plants with higher (lower) than average prices will appear to be more (less) productive than they really are. As I argue above, the geographic segmentation present in the ready-mixed industry all but ensures that within-industry demand variation exists.[17]

I measure productivity using a standard total factor productivity index. Plant TFP is computed as the log of its physical output minus a weighted sum of its logged labor, capital, materials, and energy inputs. That is,

$$TFP_{it} = q_{it} - \alpha_{lt} l_{it} - \alpha_{kt} k_{it} - \alpha_{mt} m_{it} - \alpha_{et} e_{it}, \quad (15)$$

where the weights α_j are the input elasticities of input $j \in \{l, k, m, e\}$. While inputs are plant-specific, I use the industry-level input cost shares as my measure of the corresponding input elasticities.[18] These cost shares are computed using reported industry-level labor, materials, and energy expenditures from the CM. Capital expenditures are constructed as the reported industry equipment and building stocks multiplied by their respective capital rental rates in ready-mixed concrete's corresponding two-digit industry.[19]

Labor inputs are measured as the sum of production worker hours (reported directly in the CM microdata) and an imputed value for non-production worker hours. This imputation uses the method of Davis and Haltiwanger (1991), who multiply the number of non-production

---

[17] I check below the robustness of the empirical results to the use of revenue-based output and find that some results do change from the physical-output-based benchmark, likely due to the nature of intra-industry price variation. Readers may notice that one can construct unit prices using the product-specific revenue and physical output data. I do so and explore the connection between product substitutability, costs, and prices in Syverson (2002).

[18] Two potentially important assumptions are implicit in this index. First is the assumption of constant returns to scale. If the scale elasticity were instead different from one, each of the input elasticities α_j should be multiplied by the scale elasticity. Below, I test the results for robustness to this assumption as well as estimate the scale elasticity directly in an industry production function (and, in fact, find evidence for constant returns). The second assumption is that all industry producers share the same production function—hence the use of industry-level cost shares. While this is a common assumption in similar studies, one might be concerned in particular that many producers mine intermediate materials (gravel and stone aggregate, specifically) on the factory site. Plants that do so will have quite different factor cost shares than those purchasing these materials from off-site producers. While I do not directly observe plants' sources of intermediate materials, the vast majority of plants have materials' revenue shares that are narrowly distributed around the industry average of roughly 65 percent. Given that gravel and stone are major intermediate inputs (accounting for roughly 13 percent of revenues), the narrow distribution suggests on-site mining is the exception rather than the rule.

[19] Capital rental rates are from unpublished data constructed by the Bureau of Labor Statistics for use in computing their Multifactor Productivity series. Formulas, related methodology, and data sources are described in U.S. Bureau of Labor Statistics (1983) and Harper, Berndt, and Wood (1989).

workers at the plant by the average annual hours of non-production employees in the corresponding two-digit industry (computed from Current Population Survey data). Equipment and building capital stocks are plants' reported book values of each capital type deflated by the book-to-real value ratio for the corresponding three-digit industry. (These industry-level equipment and structures stocks are from published Bureau of Economic Analysis data.) Any reported machinery or building rentals by the plant are inflated to stocks by dividing by the type-specific rental cost from Bureau of Labor Statistics data (see footnote 20). The total productive capital stock k_{it} is then computed by summing the equipment and structures stocks. Materials and energy inputs are simply plants' reported expenditures on each divided by their respective industry-level deflators from the National Bureau of Economic Research Productivity Database.[20]

B. Local Markets in the Ready-Mixed Concrete Industry

The empirical test of the above model works off variations in demand density across geographic markets. This of course raises the issue of how to suitably define markets within the industry. I use the Bureau of Economic Analysis's Component Economic Area (CEA) as my market definition. CEAs are collections of counties usually—but not always—centered on Metropolitan Statistical Areas (MSAs). Counties are selected for inclusion in a given CEA based upon their MSA status, worker commuting patterns, and newspaper circulation patterns, subject to the condition that CEAs must contain only contiguous counties. The selection criteria ensure that counties in a given CEA are economically intertwined. This classification process groups the roughly 3200 U.S. counties into 348 markets that are mutually exclusive and exhaustive of the land mass of the United States.[21]

[20] The fact that plants can produce multiple products—though most do not, as discussed above—poses another productivity measurement issue. Inputs are reported on an establishment-wide rather than product-specific basis. Thus the TFP value for multi-product concrete plants must impute the share of inputs allocated to ready-mixed production. I do so by dividing reported ready-mixed output by its share of total establishment sales. This adjustment method in effect assumes inputs are used proportionately to each product's revenue share. For example, a plant producing 1000 cubic yards of ready-mixed concrete accounting for 80 percent of its revenues will have the same TFP as a completely specialized plant producing 1250 cubic yards with same measured inputs. Without adjusting the first plant's output, it would appear less productive because the inputs it uses to make its other products would instead be attributed entirely to ready-mixed production. Again, the impact of any mismeasurement induced by this approximation is minimized by the fact that industry plants produce ready-mixed concrete almost exclusively.

[21] See U. S. Bureau of Economic Analysis (1995) for more detailed information about CEA creation.

The CEA-based market is a compromise between conflicting requirements. The model assumes that concrete markets are isolated geographic units; plants in one market competitively interact only with other plants in the same local market. Interactions with ready-mixed producers in other markets are assumed away. While there are bound to be some cross-CEA concrete sales in reality, the high transport costs of the industry make this unlikely. Detailed industry-level shipments data from the 1977 Commodity Transportation Survey support this. Ready-mixed plants shipped 94.4 percent (by weight) of their total output less than 100 miles. Discussions with industry managers also offer anecdotal evidence along these lines; stated maximum ideal delivery distances were between 30- and 45-minute drives from the plant. CEAs are large enough to minimize cross-market shipments. (An additional factor minimizing cross-market shipments is that most CEA boundaries are in outlying parts of urban areas and are thus less likely to be near areas heavily populated with concrete plants.) CEAs are also not required to adhere to state boundaries, which would sometimes place unwarranted market boundaries in economically interconnected areas. Of course, I do not want to make markets so large that there is very little competitive interaction between many of the included establishments. Plants placed in too large a market may not all respond to the same market forces—either external or the actions of industry competitors. CEAs are a suitable compromise to resolve the tension between isolating markets yet ensuring the producers within them are interconnected.

C. Demand Density

The key exogenous variable in the model is demand density. To measure this empirically, I use the log of the number of construction-sector workers per square mile in the CEA-year market. Construction sector employment is obtained from County Business Patterns data aggregated at the CEA level.[22] Land areas are from the City and County Data Book.

Construction sector employment is a suitable demand measure because the sector buys

[22] County Business Patterns data occasionally have missing observations due to data disclosure regulations. This is a small matter in the case of the construction sector (SICs 15-17), however. The sector's ubiquity and abundance of small firms allows full disclosure of total employment in nearly all counties (employment data is withheld in roughly 1.5 percent of the county-year observations in my sample). I impute employment when missing by multiplying the number of establishments in each of nine employment ranges (which are always reported) by the midpoint of their respective employment ranges, and summing the result. The impact of using imputes is likely to be even less than their proportion indicates, as the typically small nondisclosure counties are less likely to contain non-AR ready-mixed plants.

most of the ready-mixed concrete industry's output (97.2 percent, according to the 1987 Benchmark Input-Output tables). I also contend that this measure is empirically exogenous to the nature of competition among local ready-mixed concrete plants. This is because construction projects require output from a wide array of industries, so the cost share of ready-mixed alone is small. Looking at 1987 again, ready-mixed concrete accounted for only 2.0 percent of the construction sector's costs. Therefore a shock to the competitiveness of the local ready-mixed industry (that lowers average concrete prices, say) is unlikely to spur a construction boom. Thus causation travels from construction demand to concrete competitiveness and not in the reverse direction.

IV. Benchmark Empirical Results

To recap the discussion of Section II, the model implies that higher demand densities should result in the following:

- Less productivity dispersion among local producers.
- A higher average productivity level.
- A higher minimum productivity level.
- Larger average plant size in terms of units sold.

I compute six market-level measures to test these implications. These measures have been selected to account for specific measurement concerns. I measure productivity dispersion using the interquartile TFP range among producers in each CEA-year market. I use this ordinal dispersion measure to minimize the influence of spurious outliers.[23] The central tendency of the local productivity distribution is measured in two ways. One is the median TFP level in the market. Again, an ordinal measure is used to minimize measurement error. The second is a weighted-average TFP in the market. The weights are producers' output market shares (among only those plants in my sample—AR plants are excluded since I observe neither their output nor productivity level). This measure is more vulnerable to the influence of outliers, of course, but it captures productivity growth resulting from density-driven market share reallocations. As a measure of the market's minimum productivity level, which outliers affect above all other

[23] It is not uncommon for reporting and transcription error create nonsensical observations in producer-level microdata. Additionally, since many of my markets have a small number of plants, this increases the vulnerability of traditionally calculated moments to outlier effects.

moments, I use the 10th-percentile TFP plant in the local market. While for many markets this is the minimum productivity level, its use does avoid some of the more questionable bottom-end productivity levels in large markets.

Two measures of average plant size are used: the average logged output (measured in cubic yards) among market plants, and the market's producer-to-demand ratio. This is simply the number of market plants divided by total construction sector employment in the CEA-year. Since the total number of plants includes those with imputed data that are otherwise excluded from the other measures above, this latter size measure captures any influence from these smaller producers.

The use of Component Economic Areas to define geographic markets offers a potential number of 1044 observations (348 CEAs x 3 CM years). In the benchmark results, I use only the 665 CEA-year observations with at least five plants for which I observe TFP, in order to improve moment-measurement accuracy. I will test the results for robustness to this cutoff.

The empirical specification used to test for the impact of demand density on moments of the local productivity and size distributions is

$$y_{it} = \beta_0 + \beta_d dens_{it} + X_{c,it} B_c + \varepsilon_{it} . \quad (16)$$

The dependent variable (one of the six measures discussed above) in CEA *i*, year *t* is a function of local demand density $dens_{it}$, a vector $X_{c,it}$ of other influences on the moments, and an CEA-year-specific error term. I estimate three versions of this general model. A simple bivariate regression of the dependent variable on demand density characterizes the nature of the correlation between these variables. I then estimate a specification where $X_{c,it}$ contains year fixed effects. Finally, the model is estimated with $X_{c,it}$ including not only year dummies but also a number of local demand influences.

The local demand controls in $X_{c,it}$ include an assortment of variables that plausibly shift, directly or indirectly, the demand structure of the local ready-mixed concrete market through channels other than the spatial substitutability influence of demand density.[24] These include

[24] Of course, omission of these other influences will not necessary bias the demand density results in the sparser specifications if they are orthogonal to my demand density measure. It seems possible, however, that some of these control variables independently influence both local ready-mixed market structure and the overall level of construction activity (which is a key element of the demand density measure). Due to data limitations, some of these measures are CEA-specific but not CEA-year-specific. In these cases, I have attempted to use values gathered as close to the middle of the sample period as possible.

demographics of the CEA: the percentage of the population that is nonwhite, the fraction over 25 years old, the proportion with at least a bachelor's degree, and the number of marriages per 1000 population. Each of these variables is aggregated from values in the 1988 City and County Data Book. The race and the marriage variables constructed from 1984 data, while the others are from the 1980 population census. I also include variables conceivably correlated with concrete demand specifically. These are the fraction of households owning at least two automobiles, the fraction of housing units that are owner-occupied, the median value of owner-occupied housing, and median household income (also from 1980 and 1984). The output-weighted average ready-mixed specialization ratio (i.e., the fraction of plant revenue from ready-mixed) is also included to control for any systematic differences in specialization across markets. Finally, I add the growth rate of local construction employment over the previous five years to control for short-term effects (for example, a temporary boom might allow relatively inefficient producers to operate for a short while).

Panel A of Table 1 shows summary statistics of key regression variables. It is readily apparent that there are nontrivial differences in productivity moments and average plant sizes across local markets. For example, the standard deviation across markets in median productivity levels is roughly 12 percent. It is 19 percent for output-weighted average TFP, and over 50 percent for the 10^{th}-percentile productivity level. There is also variation in the amount of within-market productivity dispersion. The standard deviation of the interquartile productivity range is half of its mean of 0.274.[25] The standard deviation of average plant size (in terms of total cubic yards shipped) is over 70 percent. Panel B reports the distribution of the number of producers across all markets and for those in the sample.

The benchmark regression results are presented in Table 2. Panel A of the table shows, for each specification and dependent variable, the estimated demand density coefficients and heteroskedasticity-robust standard errors. Columns 1 through 3 of the table correspond to the specifications discussed above. I do not report covariate estimates in the interest of parsimony; Panel B instead summarizes the nature of the covariate coefficients.

The results support the predictions of the model. Productivity dispersion declines with density. The median productivity, the quantity-weighted average productivity, and the 10^{th}-

[25] An interquartile range of 0.274 in log-level TFP within a market implies that the 75^{th}-percentile productivity plant can produce roughly 30 percent more output than the 25^{th}-percentile plant with equal amounts of measured inputs.

percentile ("minimum") productivity levels are all higher in denser markets. The producer-to-demand ratio falls, and average plant output climbs with demand density. The coefficients are of the expected sign and statistically significant for every dependent variable in each of model's specifications. Beyond statistical significance, the estimates imply what are in my opinion nontrivial economic impacts. Controlling for other influences on demand, a one-standard-deviation increase in logged demand density implies a decrease in expected dispersion by approximately 0.042 log points—roughly one-seventh of the mean dispersion and over one-fourth of its standard deviation (see Table 1). The same density increase corresponds to about a 2.2 percent, 2.3 percent, and 9.4 percent increases in median, output-weighted average, and 10^{th}-percentile TFP levels, respectively. Given that demand density ranges several orders of magnitude across the markets in my sample, these imply noticeable differences in productivity levels when comparing very dense urban markets to their rural counterparts. The plant size effects implied by the regressions are quite large. For example, an average-sized plant in a market that is one standard deviation denser than another sells 30 percent more cubic yards of ready-mixed concrete than its counterpart in the lower density market.

As can be seen in the column 3, Adding market demand controls to the regression does change some of the coefficients' magnitudes. The estimated downward effect on dispersion and the upward impact on the lower-bound TFP level become even greater when local demand conditions are accounted for. On the other hand, density's estimated impact on the central tendency of the productivity distribution diminishes. The plant size effects are also of lower magnitude, but the differences are relatively small.

In the fourth column, I report the results of a specification where an additional control of special interest has been added to the $X_{c,it}$ vector. Ciccone and Hall (1996) explored the effect of market density on productivity levels. While in some ways similar in spirit to this study, their research employs a much more top-down approach, using highly aggregated production data. As such, they investigate productivity effects averaged across many industries rather than in specific sectors, and they are unable to examine differences in productivity dispersion. Further, they use an overall employment density measure to capture agglomeration effects of unspecified origin(s). This is in contrast to my industry-specific demand density measure which embodies a specific mechanism through which market density acts. (Indeed, it is possible that the mechanisms modeled and tested here are in part driving Ciccone and Hall's results.) To see if the present findings can be distinguished from their results, though, I estimate a specification that includes a measure of local employment density (using 1986 civilian employment numbers from the City and County Data Book) constructed in a similar manner to theirs.

The results indicate that the inclusion of overall density does not affect the estimated effects of demand density on the dispersion or lower bound of the local productivity distribution, or on the measures of plant size. However, demand density's influence on the median and output-weighted average productivity levels becomes small and insignificant. Clearly then the impacts of demand density and overall density on *average* productivity levels are closely related observationally. This may be because the impact of demand density (more precisely, the component independent of overall thick-market effects) is on productivity dispersion rather than levels. Alternatively, it may be that the demand density mechanism posited here is in part driving the Ciccone-Hall results, and that an overall market density measure captures the causal influence of demand density. Regardless of the specific mechanism, it is clear that ready-mixed producers in denser markets are more productive on average.

Interestingly, it seems that the transport-cost-driven substitutability mechanism I exposit explains only a modest portion of the differences in local productivity distribution moments across markets. The R^2 for the bivariate regressions in the benchmark specification indicate that demand density differences alone account for roughly 2 percent of the across-market variation in productivity dispersion. The ability of density to explain median productivity levels is stronger—6 percent—but still moderate. These values are somewhat surprising, given the perceived level of homogeneity in ready-mixed output. I will return to this issue below. Density differences do, however, explain a considerable portion of the variation in average plant sizes across markets.

V. Robustness Checks

The benchmark results are consistent with the predictions of the model. Here I conduct several checks to see if those results are robust to various empirical modeling assumptions made above. To keep things from becoming too cluttered, I only report results for the bivariate and full-demand control models (corresponding to the specifications columns 1 and 3 of Table 2, Panel A). The bivariate and full-control results of these checks are respectively presented in Tables 3 and 4. I show for comparison the benchmark estimates in column 1 of each table. A key to the different specifications is in Table 5.

A. Minimum Number of Establishments

For inclusion in my benchmark sample, I require that a market has at least five ready-mixed producers with non-imputed productivity and output data. Here I check to see if changing this exclusion criterion affects the results. Columns 2 and 3 of Tables 3 and 4 show the demand density coefficients obtained when the cutoff levels for the minimum number of plants are instead (respectively) 2 and 10. Obviously, the number of market-year observations is larger in the former case (974 markets meet the looser requirement) and smaller in the latter (343 markets). Under either alternative exclusion condition, the results largely coincide with the benchmark results. There is very little difference—qualitatively or quantitatively—between the benchmark estimates and those from the 2-plant-minimum sample. There are a few noticeable differences in the 10-plant-minimum case. The density coefficients in the productivity-moment bivariate regressions in Table 3 are smaller in magnitude than their benchmark counterparts, but they all remain statistically significant. In the model with demand controls included, however, all become insignificant except for the coefficient in the 10th-percentile productivity level regression (all do retain the expected sign).[26] Density's estimated effect on plant size is not affected in either specification using the large-market sample. Thus the results on balance largely echo the benchmark findings.

B. Capital Measurement

I exclude Administrative Record (AR) establishments from my sample to obtain TFP measures only for plants without imputed production data. There is one exception to this imputation exclusion, however: even among non-AR producers, capital stocks are imputed for those not in the Annual Survey of Manufactures (ASM) panel corresponding to that CM year

---

[26] The smaller and less precise demand density coefficients in the large-market sample could result simply because the sample is smaller and has less density variation, or alternatively because density's impact is nonlinear. To investigate this further, I estimated a specification (using the benchmark 5-minimum-plant cutoff) that included both demand density and its square. The quadratic density term was small and insignificant in the dispersion and lower-bound productivity regressions, but negative and significant in both the median and output-weighted average TFP regressions. Thus it seems that the impact of density on the central tendency of the productivity distribution fades somewhat in high-density markets. Since these markets also tend to have more producers, this may explain why I found a smaller density impact when focusing on these markets. These results are available from the author.

(these panels run for five-year periods and span a single CM). ASM plants comprise roughly one-third of my sample. Since a plant's probability of being selected for inclusion in the ASM panel increases with size, smaller concrete plants are more likely to have imputed capital stocks. Since less dense markets tend to have smaller plants—as shown above—TFP measurement error may cause productivity dispersion to be spuriously larger in low-density markets.

This potential problem is mitigated by the fact that capital cost shares are rather small in the industry (around 6.5 percent), so any spurious TFP variation due to incorrectly measured capital stocks will be small relative to the amount of measurement error. Still, to verify that capital imputations are not driving my results (particularly for productivity dispersion), I rerun the productivity-moment regressions using producers' *labor* productivity levels (the log of physical output per worker-hour) rather than TFP. I also include the market's output-weighted mean logged capital-labor ratio to account for differences in average capital intensity across markets (measurement error should be smaller for these averaged capital figures). The results, which support the benchmark findings, are shown in column 4 of Tables 3 and 4. Density coefficients are of the expected sign, and significantly so, for each of the productivity moments in both specifications. (The sizes of the coefficients are not directly comparable to the benchmarks because the dependent variables are different.) The impact of imputed capital stocks on TFP measures does not seem to be important to the qualitative features of the results.[27]

C. Output Measure

As discussed above, I measure output in physical units (cubic yards of ready-mixed concrete) to keep plant-level price variation out of my productivity measures. However, revenue-based output measures are appropriate in some instances, such as when prices embody quality differences.[28]

---

[27] Variable capital utilization poses another possible form of capital measurement error. If utilization differs substantially across producers, capital stocks do not accurately measure the amount of capital services used in production. To assess the influence of variable capital utilization on my results, I re-estimated the empirical model using plant productivity levels from an index suggested by Basu and Kimball (1997). This procedure uses hours per worker as a proxy for capital intensity by assuming that production is Leontief in value added and materials. The results of this exercise, available from the author, are not substantially different from the benchmark results.

[28] Abbott (1992) discusses this issue in considerable detail. While ready-mixed concrete is much more homogeneous than many manufactured goods, there is some scope for differentiation. By changing ingredient ratios and using admixtures, producers can achieve variation in the physical and aesthetic properties of ready-mixed concrete. If plants differ in the proportions with which they produce these various product types, and type-specific

While I cannot test for quality effects directly, I can see if the results are affected by the output measure. The results obtained when total plant revenue (deflated to 1987 dollars) is used in place of physical output to compute TFP are presented in column 5 of Tables 3 and 4. Consider first the results from the bivariate specification in Table 3. The signs and statistical significance of the benchmark results are preserved, and the magnitudes of the effects on productivity dispersion and average plant size are virtually unchanged.[29] However, density's estimated impact on the median, output-weighted average, and 10^{th}-percentile TFP levels is smaller. These central tendency and lower-bound results are weaker still in the specification that includes demand controls. The estimates in the median and output-weighted mean productivity regressions are virtually zero and insignificant (indeed, for the first time a coefficient—in the output-weighted mean TFP regression—does not have the expected sign, albeit insignificantly). The productivity dispersion and average plant size (where size is now measured in deflated revenue) coefficients remain very close to their benchmark counterparts.

The weakened estimated influence of density on the central tendency and lower bound of the productivity distribution is interesting and worth a brief discussion. Given that physical productivity grows but revenue productivity remains constant in density, average prices must be lower in dense markets.[30] If prices reflect quality differences, then average quality is lower in denser markets. This seems counterintuitive. Furthermore, plants in denser markets would expectedly be more likely to be specialized—some would focus on production of low-quality ready-mixed while others in high-quality—which should induce a positive correlation between revenue-based TFP dispersion and density. Yet I find the opposite. Thus on two levels, the results seem unlikely to be driven by product quality variation. An alternative, and in my opinion more likely explanation for the outcomes here is that prices decline in density because both costs and markups are lower in denser markets (the markup effect is due to the higher demand elasticities created by greater substitutability). This would make average revenue-based productivity measures less responsive to density changes, as found here. It is also consistent

---

prices reflect differences in material and production costs, quality differences will be reflected in plant revenues.

[29] In this specification, I measure average plant output as the log of total *revenue* per plant. Unlike the benchmark case, where I compute average plant output for only those plants for which I have production data, all producers are counted in the revenue-based size measure. Total plant sales are one of the few non-imputed variables in AR data.

[30] I show this is the case directly using computed unit prices in Syverson (2003).

with the unchanged productivity dispersion result, because cost-based price dispersion (and therefore revenue-based TFP dispersion) would be negatively related to density.

D. Scale (Dis)Economies

The TFP measures used above assume constant returns to scale in the plant-level production technology. Here I see how the benchmark results are affected by departures from this assumption. Column 6 of Tables 3 and 4 show the results obtained when the assumed scale elasticity is 0.9 instead of one (that is, all input elasticities in (15) are multiplied by this value). As can be seen, the qualitative features of the benchmark estimates are preserved. The demand density coefficients are of the expected sign and statistically significant for every productivity moment. In fact, the estimated effects of demand density on the central tendency and lower bounds of market productivity distributions are larger than the benchmark.

Column 7 of the tables show the corresponding estimates when the scale elasticity is 1.1. The lower-bound and dispersion of productivity effects remain similar to the benchmark. Here, however, the magnitude and statistical significance of density's impact on the median and output-weighted mean productivity levels have fallen to near zero. A careful consideration of how the TFP index is calculated and the cross-sectional features of the industry provide a likely explanation for these results. As the assumed scale elasticity grows, increases in output that coincide with input growth are attributed increasingly to scale economies rather size-neutral productivity gains. Thus if larger industry plants tend to be more efficient because of an external selection mechanism like the impact of greater output substitutability in denser markets, assuming economies of scale in the TFP measure will attribute this size-productivity correlation to scale effects instead. This lowers the productivity estimates of larger establishments relative to smaller ones, shrinking the demand density coefficient and possibly even creating a negative correlation if the assumed scale elasticity is sufficiently large.[31] Notice that this would result even if the industry technology's true *internal* scale elasticity was one.

The upshot, then, is that the estimated effects of demand density on productivity dispersion and average plant size appear robust to the assumed scale elasticity of the industry's technology, but the impact on the average productivity level is more sensitive to the assumed

[31] This mechanism operating in the other direction likely explains the larger estimated density coefficients in the average productivity level regressions when the assumed scale elasticity is 0.9.

degree of returns to scale. If the technology exhibits diseconomies of scale, the connection strengthens beyond that found in the benchmark. Increasing returns to scale, on the other hand, weaken it. I show below, however, that the benchmark assumption of constant returns to scale seems justified.

E. Production Function Estimation

Index-number methods are a computationally simple way to compute TFP levels, but they have their own well known set of shortcomings. An alternative methodology is to back out producer TFP levels as the residuals of an estimated production function. However, the naive procedure of simply regressing output on a functional form of inputs using ordinary least squares could lead to biased production function and TFP estimates due to the "transmission bias" of productivity levels affecting input choices (first pointed out by Marschak and Andrews 1944). Much of the recent literature has tried to circumvent this endogeneity problem through the use of semiparametric methods (originally proposed by Olley and Pakes 1996, and modified by Levinsohn and Petrin 2003). Unfortunately, as discussed extensively in Syverson (2001), industries with local demand markets such as ready-mixed concrete are a particularly poor fit for such methods. So I instead take advantage of the local nature of the industry to identify instrumental variables for use in obtaining consistent production function and TFP estimates. I have already argued that local demand—as measured by the size of the concrete sector in the market—is both relevant to the amount of ready-mixed concrete produced in the market and exogenous to shifts in concrete plants' productivity levels. Therefore, measures of local demand (construction sector employees—here in levels rather than in a density measure) should be correlated with concrete plants' inputs but orthogonal to movements in their productivity levels, making them good instruments for endogenous inputs in the production function.[32]

To estimate plant TFP levels, I estimate the following production function for the ready-mixed plants in my sample:

[32] Shea (1993) discusses how input-output patterns in production can be used to identify industry-level demand shifters. I extend this insight for use with microdata by measuring and matching local downstream demand to upstream producers in the same market. In application, rather than use local construction employment in my instruments, I use the component that is orthogonal to overall local employment (obtained as the residuals of a regression of construction on overall employment in all market-year observations). This is because market-level productivity shocks common to local producers in all industries could lead to ready-mixed plants' productivity levels being correlated with local construction activity, despite concrete's small cost share in construction.

$$q_{it} = \gamma_o + \beta_m + \delta_t + \gamma_{mult} d_{mult,it} + \gamma_x x_{it} + \omega_{it}, \text{ where } x_{it} = \alpha_{lt} l_{it} + \alpha_{kt} k_{it} + \alpha_{mt} m_{it} + \alpha_{et} e_{it}.$$

As with the TFP index in (15), output and inputs are measured in logs and the input elasticities α_{jt} are measured as the industry-level cost shares. The production function includes a year effect δ_t as well as a multi-plant dummy $d_{mult,it}$ that captures any effect of operating as part of a multiple-establishment firm. A cost-share-weighted sum of logged labor, capital, materials, and energy inputs comprise the producer's composite input x_{it}. Thus γ_x is the scale elasticity of the industry technology.[33] A CEA fixed productivity effect β_m is included because, as discussed above, the model implies an across-market form of increasing returns: plants are both larger and more productive on average in denser markets. However, this is driven by selection rather than any internal scale economies within plants. So as not to confound this across-market "scale elasticity" with my internal returns to scale estimate, then, I identify γ_x using only within-market variation. I do include these fixed effects in producers' TFP levels, though, because I want to measure those systematic across-market differences in average productivity levels. Hence the TFP estimate for a producer is equal to the sum of the estimates of its corresponding market effect β_m and its plant-year-specific residual ω_{it}. Given the "transmission bias" described above, I instrument for the plant composite input x_{it} using the current value, three lags, and one lead of the local construction sector activity measure discussed above.[34]

Table A.1 shows the estimated parameters of the production function. The first-stage relevance statistics indicate that local construction activity is germane to concrete plants' hiring of inputs, even after removing the influence of overall local economic activity. The F-test for joint significance of the instruments (the three lags, present, and one lead of local construction activity) soundly rejects the null of irrelevance. Higher downstream demand increases input use, as one would expect; the sum of the instrument coefficients in the first-stage regression is positive and significant. As for the production function estimates, the scale elasticity γ_x is precisely estimated and statistically indistinguishable from one. Thus is appears that assumption

[33] I discuss in Syverson (2001) why a composite input is used rather than including each input in the production function separately.

[34] This lag/lead pattern was chosen based on two considerations. The first is my prior belief about the extent of management decision horizons, both forward- and backward-looking. The second consideration is Buse's (1992) demonstration that superfluous instruments in an instrument set lead to estimation biases. Notice that the instruments are not plant specific, but rather vary across markets and years. Given the market fixed effects in production, identification comes off of plants' responses to shifts in market demand over time.

above of constant (internal) returns to scale is justified.

Given that the benchmark results were obtained imposing a scale elasticity very close to the value estimated in the production function, it is not surprising that the results using estimated productivity levels closely match those obtained with a TFP index above. In fact, the estimated and index TFP series have a correlation of 0.984. The productivity and size moment regression results are shown in Column 8 of Tables 3 and 4. There is no substantial differentiation between the estimated-TFP and benchmark results in the bivariate regressions. In the specifications that include demand controls, density's implied effects on measures of the central tendency of TFP are slightly smaller than in the benchmark case. As a result, these estimates become marginally insignificant (both have p-values below 0.06). Excepting this, the estimated-TFP results mirror the benchmark results.

F. Using Intertemporally-Averaged Productivity Estimates

The steady-state model employed earlier implies long-run, cross-sectional differences in local productivity and size distributions across markets of varying densities. My empirical tests use cross-sectional panels at five-year intervals and control for short-term density changes in order to isolate long-term impact of demand density. As a further check against any confounding effect of short-run fluctuations, I estimate the demand density regressions using the average productivity level of each producer during the sample. That is, I compute the plant's TFP index in each of the years I observe it and use their average when computing moments of the local productivity distribution. Thus in this case there is only one cross-section of markets rather than three. All explanatory variables, including demand density, are averaged over time as well (they are weighted by the number of plants observed in each CEA-year market).

The results of these exercises can be found in column 9 of the tables of robustness checks. As can be seen, the benchmark findings are not unduly influenced by short term fluctuations. The demand density coefficients retain their expected signs and significance. Furthermore, their magnitudes closely correspond to the benchmark, with the exception of the smaller estimates in the average plant output regressions.

G. Density or Size?

The intuitive argument sketched in the introduction implies that market density specifically, rather than just market size, drives the truncation of the productivity distribution. While they are very likely to be correlated, density has a direct impact on spatial substitutability while size alone does not. I test whether the size-density distinction holds empirically with a specification that includes both demand density and demand size (logged construction employment in the local market). The results are shown in column 10 of Tables 3 and 4.

The demand density coefficients retain their expected signs in all cases, and are statistically significant in all cases, excepting the output-weighted average productivity regression including demand controls. Their magnitudes do decline somewhat (except for those in the productivity dispersion regressions, which rise insignificantly); the drops are most noticeable in the average size regressions. As for the coefficients on market size (not shown), they are significant in every bivariate-model case excepting the dispersion regression. However, once the other demand controls are included, the median and output-weighted means TFP size coefficients also lose their significance. This is perhaps not surprising given several components of $X_{c,it}$ are controls for shifts in demand.

Thus demand density has a separately identifiable influence on the local productivity and size distributions from that of market size alone. The same cannot always be said of market size. This supports the mechanism exposited in the model. Still, the two effects do overlap some, as most of the estimated density effects become smaller in magnitude when size is included in the productivity and size moment regressions.

VI. Conclusion

This paper has modeled and offered supporting evidence for a mechanism where a demand-side market characteristic—product substitutability, specifically—affects the shape of the equilibrium distribution of producers' productivity levels. The essence of the mechanism is that heightened competition (created by greater spatial substitution possibilities) in denser markets makes it harder for inefficient producers to profitably operate. This truncates the lower end of the market productivity distribution, leading to increases in its lower bound and central tendency and decreases in the amount of within-market dispersion. It further implies that producers are larger on average in denser markets. In a case study of the ready-mixed concrete industry, I find robust empirical support for each of these features.

These findings of this case study have several implications. Most directly, they suggest that spatial substitution barriers may account for some of the persistent within-industry productivity dispersion documented by other studies. More broadly, the paper's findings suggest that other factors that limit product substitutability, such as physical product differentiation, also support the persistent productivity differences observed in most industries. Syverson (2004) extends the work of this paper (with some loss in the ability to control for technological differences across producers) to explore these impacts at the industry level.

The results also support the notion that competition spurs efficiency in production. This is a comfortable concept for many economists (although a substantial portion of the long and rich literature on the topic focuses on conditions in which competition may not be efficiency-inducing). Still, there has been a relative dearth in empirical studies that have been able to carefully measure the productivity response to exogenous differences in market competitiveness. This paper has hopefully reduced that comparative scarcity.

The paper further shows how local competition in markets where transport costs are important creates an agglomeration effect (i.e., producers are more efficient in denser markets) that does not appeal to the technological externalities often discussed in the literature. Lowering transport costs thus boosts productivity in two ways. Fewer resources are spent on moving goods around, of course, but competition also increases within spatially differentiated industries, reallocating market shares to more efficient producers. This mechanism has implications ranging from the urban to trade literatures, and may be an interesting area for further research.

Still, the empirical evidence also suggests that much work remains to be done to completely characterize the nature and sources of persistent productivity differences across producers. Even in the "controlled" environment of an industry case study, observable substitutability factors still only account for a small fraction of the observed across-market variance of productivity moments. The supply-side factors that have been a primary focus of the literature to this point doubtlessly account for some of this. Perhaps as well, though, the results above suggest a significant role for unmeasured (and in many cases, unmeasurable) product differentiation—subtle variations in product attributes, subjective product differentiation like brand effects, and dissimilarities in bundled goods—in explaining why we see such stark efficiency differences across plants.

Appendix

A. Ensuring All Producers Sell Positive Quantities

To ensure that all entrants choosing to commence production sell positive quantities, the ex-post realized sales for any producer i in the area of the market between itself and its neighbor j must be greater than zero. That is, in a market where n firms have entered and the producer pair have set prices p_i and p_j,

$$x_{i,j} = \frac{p_j - p_i}{2t} + \frac{1}{2n} > 0 \,. \qquad \text{(A.1a)}$$

This condition must hold for every pair (i, j) of producers.

Using the optimal pricing equation (4), we can rewrite this in terms of the cost difference between producers i and j.

$$x_{i,j} = \frac{c_j - c_i}{4t} + \frac{1}{2n} > 0 \,. \qquad \text{(A.1b)}$$

Note that this condition will always be met if it holds when $c_j - c_i$ equals the maximum possible difference bewteen producers; that is, when $c_i = c^*$ and $c_j = 0$. In this case, (A.1b) can be rewritten as

$$c^* < \frac{2t}{n} \,. \qquad \text{(A.2)}$$

From the condition that the marginal producer expects to earn zero operating profits, as embodied in equation (7), it must be that

$$2tE\left(\frac{1}{n}\right) = c^* - E(c) + \sqrt{\frac{4tf}{D}} \,. \qquad \text{(A.3)}$$

This can be used along with (A.2) to show that

$$c^* < \frac{2t}{n} < 2tE\left(\frac{1}{n}\right) = c^* - E(c) + \sqrt{\frac{4tf}{D}} \,, \qquad \text{(A.4)}$$

because $1/n$ is a convex function.

Therefore, (A.4) shows that to ensure that all producers sell positive quantities,

$$E(c) < \sqrt{\frac{4tf}{D}} \,. \qquad \text{(A.5)}$$

Solving explicitly for $E(c)$, which is itself a fuction of c^*, in terms of parameters depends of course on the particular cost distribution $g(c)$. But notice that dc^*/df is negative:

$$\frac{\partial V^e}{\partial f} = \int_0^{c^*} \left[\frac{D}{2t} \left(c^* - c + \sqrt{\frac{4tf}{D}} \right) \left(\frac{1}{2} \sqrt{\frac{4t}{Df}} \right) - 1 \right] g(c)\,dc \,, \qquad \text{(A.6a)}$$

which simplifies to

$$\frac{\partial V^e}{\partial f} = \int_0^{c^*} \frac{1}{2} \sqrt{\frac{D}{tf}} (c^* - c) g(c)\,dc > 0 \,, \qquad \text{(A.6b)}$$

and thus $dc^*/df < 0$ by applying the implicit function theorem to (10). This means that $dE(c)/df < 0$ as well. Since the left-hand-side of (A.5) decreases in f and the right-hand-side increases, sufficiently large values of f (for fixed values of the other parameters) will ensure that the condtion holds.

B. Computing Upper-Bound Costs for "Ex-Post" Entrants

I computationally simulate a version of the model where an "ex-post" entrant is allowed to enter the market after having observed rivals' cost levels. The first step involves solving the equilibrium from the benchmark asymmetric information model to determine c^* and $E(1/n)$ for the market. For simplicity, marginal costs are assumed to be uniformly distributed on [0,1], $t = 0.1$, $s = 0.01$, and $f = 1$. (This value of f is large enough to ensure that, for the range of density levels used in these simulations, all equilibrium producers sell positive quantities. See Section A of this appendix, above.) The equilibrium solution provides the number of producers that commence production (i.e., the number of cost draws taken), which is equal to the smallest integer that is greater than the equilibrium value of $[E(1/n)]^{-1}$. It also pins down the upper bound of the uniform distribution over which producers' costs are drawn.

The equilibrium producers are placed randomly around the circle, and then for each location between market producers, (14) is set equal to zero and inverted to solve for c_i . (Note that this equation assumes incumbents can optimally reprice against the new entrant and the other incumbents.) The highest of these values in the market is the largest possible cost level of an ex-post entrant supportable by the market structure. This process is repeated 1000 times at the same density level (but with different producer cost draws), and the average and maximum upper-bound cost across these 1000 trials are computed.

The entire simulation is then repeated for several demand density values D. While this ex-post scenario is not a formal equilibrium, it seems likely that adding more complicated strategic interactions would not result in different implications regarding the nature of the productivity truncation, particularly since this exercise makes the conservative assumptions mentioned in the text.

References

Abbott, Thomas A. "Price Dispersion in U.S. Manufacturing: Implications for the Aggregation of Products and Firms." Center for Economic Studies Working Paper no. 92-3, March 1992.

Aghion, Philippe, and Howitt, Peter. "A Model of Growth through Creative Destruction." *Econometrica*, *50*, 1992, pp. 323-357.

Asplund, Marcus and Volker Nocke. "Firm Turnover in Imperfectly Competitive Markets." PIER Working Paper 03-010, 2003.

Baily, Martin Neil, Charles Hulten, and David Campbell. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity. Microeconomics*, 1992, pp. 187-249.

Bartelsman, Eric J. and Doms, Mark. "Understanding Productivity: Lessons from Longitudinal Microdata." *Journal of Economic Literature*, *38*(3), 2000, pp. 569-95.

Basu, Susanto and Kimball, Miles S. "Cyclical Productivity with Unobserved Input Variation." NBER Working Paper no. 5915, February 1997.

Bresnahan, Timothy F. and Reiss, Peter C. "Entry and Competition in Concentrated Markets." *Journal of Political Economy*, *99*(5), 1991, pp. 997-1009.

Buse, A. "The Bias of Instrumental Variables Estimators." *Econometrica*, *60*(1), 1992, pp. 173-80.

Campbell, Jeffrey R. and Hugo A. Hopenhayn. "Market Size Matters." NBER Working Paper No. 9113, 2002.

Ciccone, Antonio and Hall, Robert E. "Productivity and the Density of Economic Activity." *American Economic Review*, *86*(1), 1996, pp. 54-70.

Davis, Steven J. and Haltiwanger, John. "Wage Dispersion Between and Within U.S. Manufacturing Plants, 1963-86." *Brookings Papers on Economic Activity: Microeconomics*, 1991, pp. 115-80.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. "The Growth and Failure of U.S. Manufacturing Plants." *Quarterly Journal of Economics*, *104*(4), 1989, pp. 671-98.

Foster, Lucia, John Haltiwanger, and C. J. Krizan. "The Link Between Aggregate and Micro Productivity Growth: Evidence from Retail Trade." NBER Working Paper 9120, 2002.

Harper, Michael J., Ernst R. Berndt, David O. Wood. "Rates of Return and Capital Aggregation Using Alternative Rental Prices." in D.W. Jorgenson and R. London, *Technology and*

Capital Formation, MIT Press, 1989.

Hopenhayn, Hugo A. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, *60*(5), 1992, pp. 1127-50.

Jovanovic, Boyan. "Selection and the Evolution of Industry." *Econometrica*, *50*(3), 1982, pp. 649-70.

Levinsohn, James and Petrin, Amil. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies*, *70*(2), 2003, pp. 317-342..

Marschak, Jacob and Andrews, William H. "Random Simultaneous Equations and the Theory of Production." *Econometrica*, *12*(3/4), 1944, pp. 143-205.

Melitz, Marc J. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* , *71*(6), 2003, pp.1695-1725.

Nickell, Stephen J. "Competition and Corporate Performance." *Journal of Political Economy*, *104*(4), 1996, pp. 724-746.

Olley, Steven G. and Pakes, Ariel. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, *64*(4), 1996, pp.1263-97.

Raith, Michael. "Competition, Risk, and Managerial Incentives." *American Economic Review*, *93*(4), 2003, pp. 1425-36.

Salop, Steven C. "Monopolistic Competition with Outside Goods." *Bell Journal of Economics*, *10*(1), 1979, pp. 141-56.

Schmitz, James A. "What Determines Labor Productivity? Lessons from the Dramatic Recovery of the U.S. and Canadian Iron-Ore Industries Since Their Early 1980s Crisis." Federal Reserve Bank of Minneapolis Research Department Staff Report 286, 2004.

Shea, John. "The Input-Output Approach to Instrument Selection." *Journal of Business and Economic Statistics*, *11*(2), 1993, pp. 145-55.

Syverson, Chad. "Market Structure and Productivity: A Concrete Example." Mimeo, University of Chicago, 2001.

Syverson, Chad. "Price Dispersion: The Role of Product Substitutability and Productivity." Mimeo, University of Chicago, 2003.

Syverson, Chad. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics*, *86*(2), 2004.

U.S. Bureau of Economic Analysis. "Redefinition of the BEA Economic Areas." *Survey of*

Current Business, February 1995, pp.75-81.

U.S. Bureau of Labor Statistics. *Trends in Multifactor Productivity, 1948-81*. Bulletin 2178, 1983.

Figure 1. Total Factor Productivity Kernel Density Estimates, Plants in Markets above and below Median Demand Density

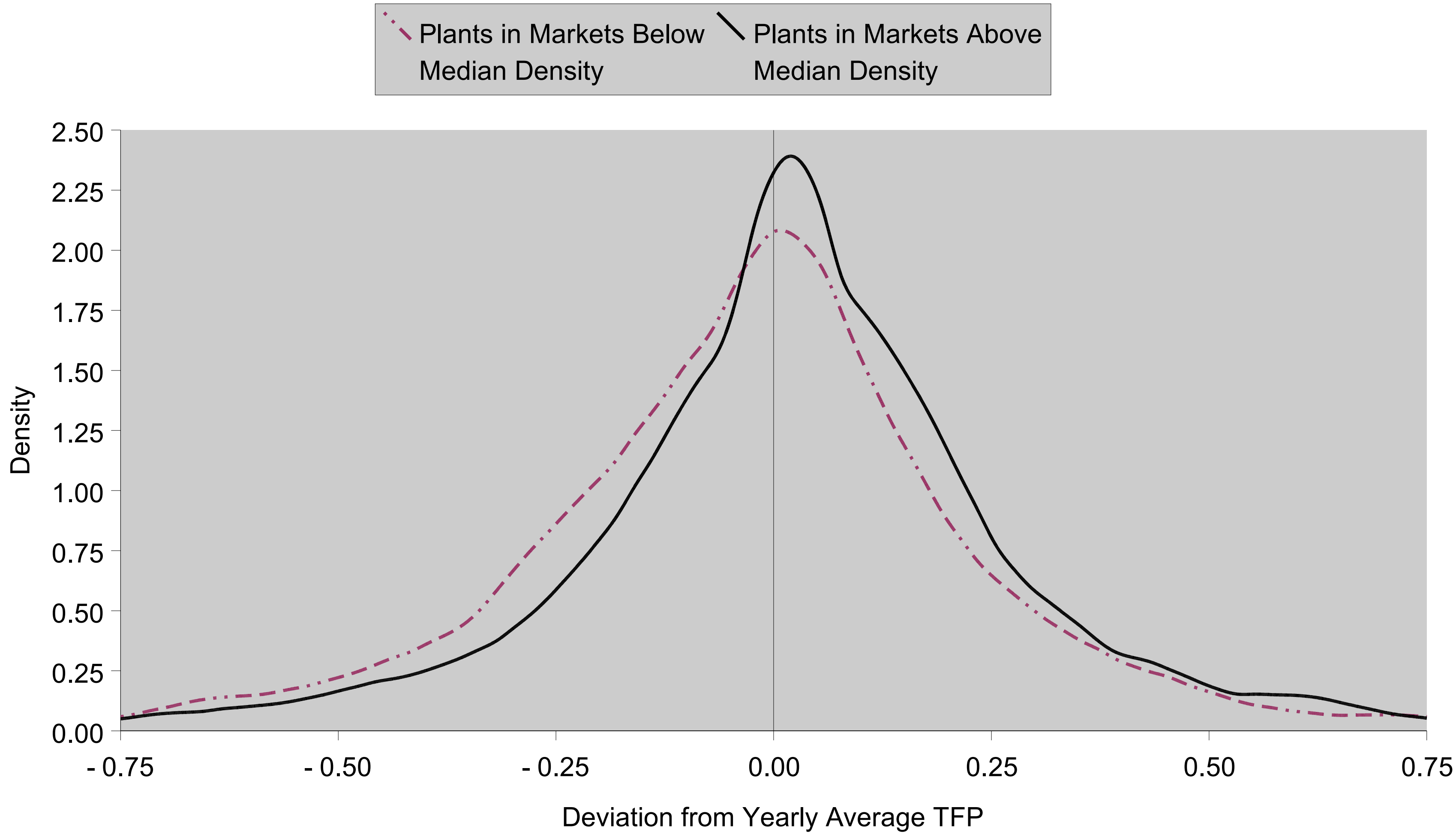

Figure 2. Output (Logged Shipments in Cubic Yards) Kernel Density Estimates, Plants in Markets above and below Median Density

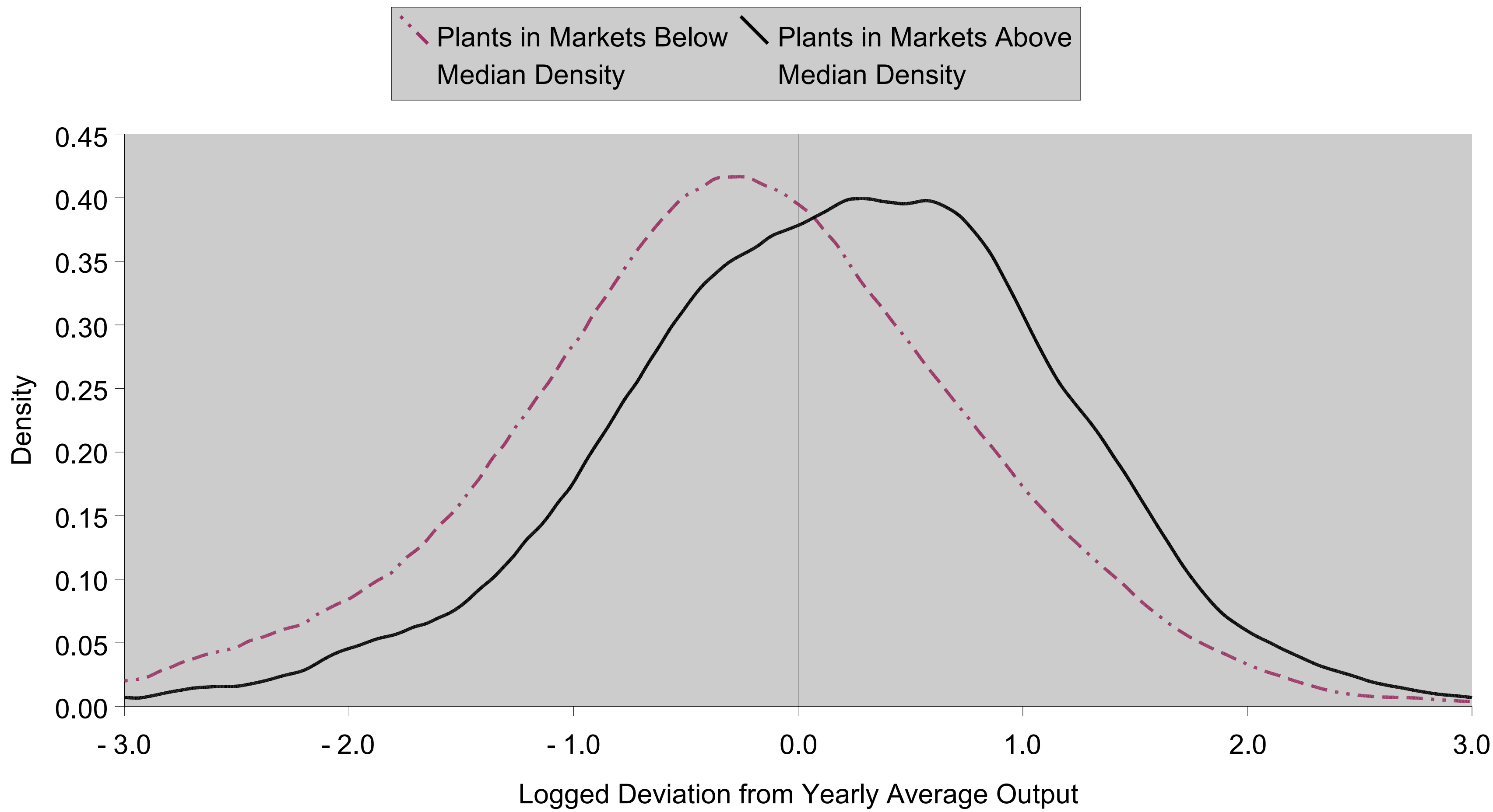

Figure 3. Simulated Upper-Bound Marginal Costs of "Ex-Post" Entrants and Average Producer Size vs. Demand Density

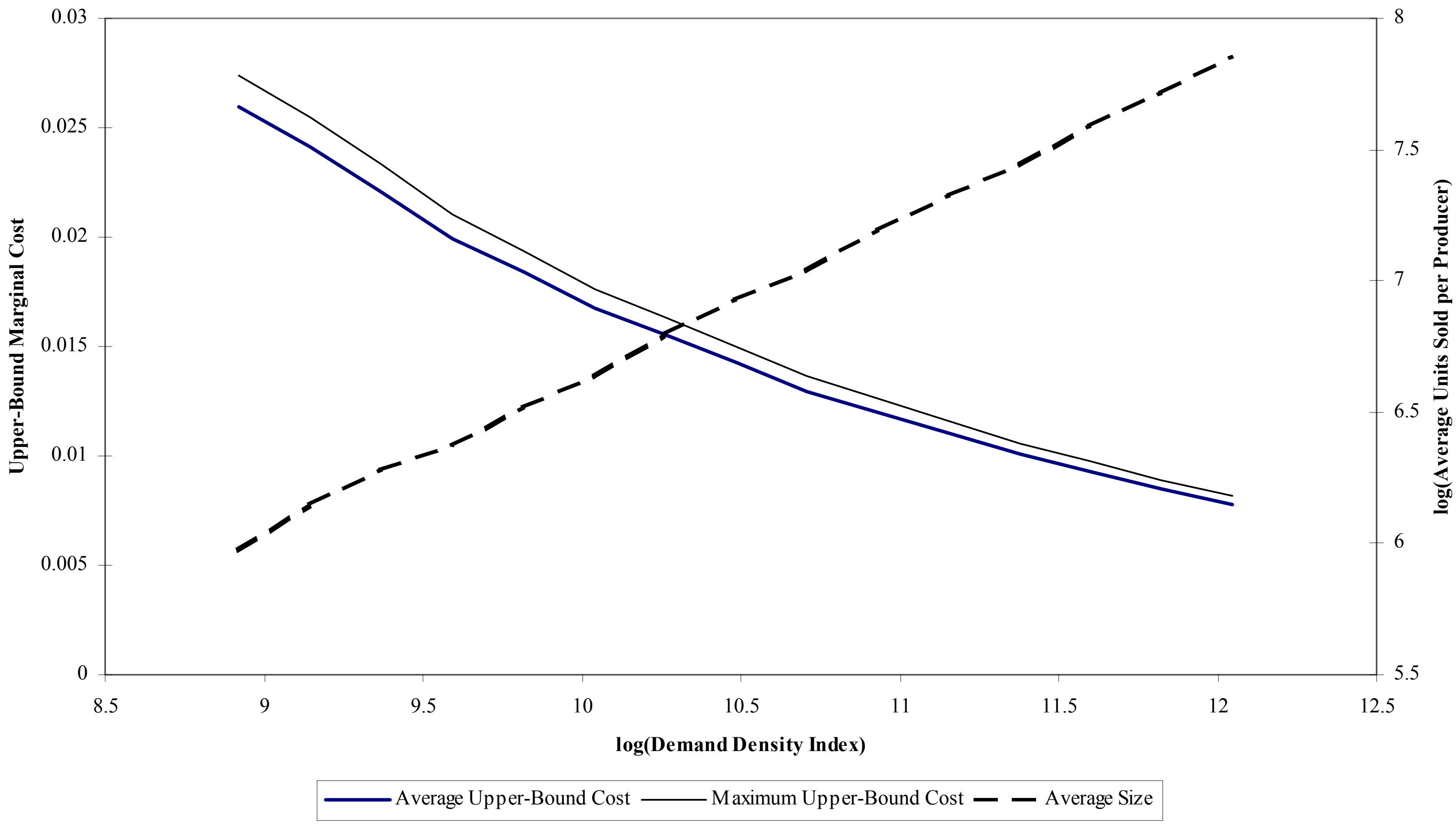

Table 1. Descriptive Statistics

A. Productivity and Size Moments and Demand Density

Variable	Mean	Std. Dev.	Skewness	75^{th}-25^{th} Percentile Range	90^{th}-10^{th} Percentile Range
TFP Dispersion (Interquartile Range)	0.275	0.147	1.332	0.164	0.353
Median TFP	4.557	0.122	-0.028	0.133	0.278
Output-Weighted Average TFP	4.599	0.177	1.627	0.183	0.360
10^{th} Percentile TFP	4.169	0.450	-6.114	0.293	0.611
ln(Plants per Demand Unit)	-6.447	0.703	-0.078	0.978	1.827
ln(Average Plant Output)	10.637	0.543	0.086	0.752	1.343
Demand Density [ln(constr. emp./mi^2)]	0.495	1.439	0.050	1.700	3.691
Demand Density—All 1044 U.S. Markets	0.317	1.407	0.066	1.534	3.548
exp(Demand Density)	4.801	11.233	6.009	3.269	10.777
ln(Number of Construction Employees)	9.229	1.035	0.401	1.547	2.690
Plant-level TFP (N = 10,613)	4.559	0.354	-1.655	0.277	0.637

B. Number of Producers by Market

Market Set	Variable	Mean	Min	25%ile	Median	75%ile	Max
All Markets w/One or More Plants (N=1031)	Number of Plants	15.0	1	6	10	18	102
	Number w/TFP Data	10.3	1	3	7	12	79
Estimating Sample	Number of Plants	20.6	5	10	14	25	102
	Number w/TFP Data	14.5	5	7	10	17	79

This table shows descriptive statistics for key variables. Figures are across 665 market-year observations unless otherwise stated. See text for details.

Table 2. Main Regression Results—Local Productivity and Size Moments

A. Demand Density Coefficients

Dependent Variable	Regression Statistic	Model [1]	Model [2]	Model [3]	Model [4]
TFP Dispersion (Interquartile Range)	R^2	0.018	0.036	0.092	0.092
	Demand Density Coef. (Standard Error)	-0.014* (0.004)	-0.015* (0.004)	-0.029* (0.008)	-0.031* (0.010)
Median TFP	R^2	0.059	0.289	0.321	0.322
	Demand Density Coef. (Standard Error)	0.021* (0.003)	0.018* (0.003)	0.012* (0.005)	0.008 (0.006)
Output-Weighted Average TFP	R^2	0.045	0.125	0.162	0.162
	Demand Density Coef. (Standard Error)	0.026* (0.004)	0.024* (0.004)	0.016* (0.008)	0.012 (0.010)
10^{th} Percentile TFP	R^2	0.033	0.033	0.058	0.059
	Demand Density Coef. (Standard Error)	0.057* (0.010)	0.056* (0.010)	0.065* (0.019)	0.056* (0.022)
Producer-to-Demand Ratio	R^2	0.570	0.584	0.708	0.711
	Demand Density Coef. (Standard Error)	-0.369* (0.015)	-0.363* (0.015)	-0.313* (0.022)	-0.278* (0.030)
Average Plant Output	R^2	0.334	0.376	0.557	0.563
	Demand Density Coef. (Standard Error)	0.218* (0.012)	0.211* (0.012)	0.184* (0.017)	0.142* (0.023)
Year Dummies		No	Yes	Yes	Yes
Demand Controls		No	No	Yes	Yes (+CH)

This panel shows the estimated coefficients on demand density when various moments of the local productivity and size distributions are regressed on demand density and, when applicable, a set of demand controls. Specifications are by column and dependent variables by row. The sample consists of 665 region-year observations with at least five plants for which I have non-imputed production data. "+CH" indicates that Ciccone-Hall measure of overall density was included in controls (see text for details). Reported standard errors are robust to heteroskedasticity, and an asterisk denotes significance at the 5 percent level.

Table 2 (continued). Main Regression Results—Local Productivity and Size Moments

B. Significance of Demand Controls

Dependent Variable	Negative and Significant	Positive and Significant
TFP Dispersion (Interquartile Range)	1982, 1987, Married, College, Auto2, Spec	MedIncome
Median TFP	1982, 1987, Nonwhite, Occupied	MedIncome
Output-Weighted Average TFP	1982, 1987, Black	MedIncome
10^{th} Percentile TFP	None	Married, College
Producer-to-Demand Ratio	Married, Nonwhite, College, MedIncome, Growth	1982, Over25, Spec
Average Plant Output	1982, Over25	1987, Married, College, MedHouse, Growth

This panel shows, by dependent variable, the significance of the demand controls included in the specification corresponding to column 3 in panel A. All demand controls are included in each regression, so those not reported were statistically insignificant.

Key to Demand Controls:
1982—Year dummy
1987—Year dummy
Married—Fraction of population that is married
Nonwhite—Fraction of population that is non-white
College—Fraction of population with college education
Over25—Fraction of population over 25 years old
MedIncome—Logged median household income
Auto2—Fraction of households with at least two cars
Occupied—Fraction of owner-occupied housing units
MedHouse—Logged value of median home
Growth—Demand growth over past 5 years (log change in construction sector employees)
Spec—Output-weighted average revenue share of ready-mixed concrete among concrete plants

Table 3. Robustness Checks—Bivariate Specification (Only Demand Density and a Constant Included in Regressions)

Dependent Variable		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TFP Dispersion	R^2	0.018	0.026	0.016	0.018	0.020	0.015	0.011	0.019	0.018	0.019
	Demand Density	-0.014*	-0.026*	-0.011*	-0.016*	-0.012*	-0.012*	-0.011*	-0.014*	-0.014*	-0.016*
	(s.e.)	(0.004)	(0.006)	(0.005)	(0.008)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
Median TFP	R^2	0.059	0.036	0.021	0.359	0.034	0.167	0.003	0.050	0.074	0.072
	Demand Density	0.021*	0.021*	0.011*	0.072*	0.013*	0.040*	0.005	0.017*	0.021*	0.015*
	(s.e.)	(0.003)	(0.003)	(0.004)	(0.007)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Qty-Wt. Average TFP	R^2	0.045	0.038	0.009	0.413	0.028	0.113	0.004	0.034	0.037	0.056
	Demand Density	0.026*	0.027*	0.012*	0.080*	0.014*	0.044*	0.008	0.022*	0.021*	0.016*
	(s.e.)	(0.004)	(0.004)	(0.005)	(0.006)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
10^{th} Percentile TFP	R^2	0.033	0.024	0.028	0.142	0.054	0.052	0.015	0.028	0.071	0.040
	Demand Density	0.057*	0.046*	0.027*	0.120*	0.027*	0.070*	0.038*	0.053*	0.064*	0.032*
	(s.e.)	(0.010)	(0.008)	(0.008)	(0.013)	(0.004)	(0.009)	(0.010)	(0.010)	(0.014)	(0.013)
Producer-to-Demand Ratio	R^2	0.570	0.510	0.576						0.207	0.713
	Demand Density	-0.369*	-0.361*	-0.365*						-0.320*	-0.212*
	(s.e.)	(0.015)	(0.012)	(0.023)						(0.033)	(0.016)
Average Plant Output	R^2	0.334	0.269	0.311		0.353				0.222	0.467
	Demand Density	0.218*	0.211*	0.206*		0.227*				0.182*	0.098*
	(s.e.)	(0.012)	(0.011)	(0.017)		(0.013)				(0.020)	(0.015)

This table shows results from robustness tests on the results presented in Table 2. See Table 5 for the key to specifications (by column). Heteroskedasticity-robust standard errors are reported. Average plant output is revenue-based in specification 5.

Table 4. Robustness Checks—Year Effects and Demand Controls Included

Dependent Variable		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TFP Dispersion	R^2	0.092	0.065	0.103	0.044	0.083	0.062	0.088	0.093	0.072	0.090
	Demand Density	-0.029*	-0.034*	-0.008	-0.031*	-0.021*	-0.024*	-0.028*	-0.033*	-0.017*	-0.029*
	(s.e.)	(0.008)	(0.009)	(0.008)	(0.012)	(0.005)	(0.008)	(0.009)	(0.009)	(0.007)	(0.009)
Median TFP	R^2	0.321	0.181	0.429	0.414	0.386	0.393	0.252	0.110	0.151	0.337
	Demand Density	0.012*	0.017*	0.006	0.054*	0.002	0.028*	0.000	0.009	0.013*	0.010*
	(s.e.)	(0.005)	(0.005)	(0.007)	(0.011)	(0.004)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)
Qty-Wt. Average TFP	R^2	0.162	0.133	0.165	0.469	0.311	0.249	0.120	0.078	0.107	0.168
	Demand Density	0.016*	0.023*	0.013	0.056*	-0.004	0.032*	-0.000	0.014	0.016*	0.011
	(s.e.)	(0.008)	(0.006)	(0.011)	(0.009)	(0.005)	(0.008)	(0.008)	(0.008)	(0.006)	(0.008)
10th Percentile TFP	R^2	0.058	0.051	0.110	0.176	0.154	0.076	0.041	0.066	0.121	0.063
	Demand Density	0.065*	0.058*	0.033*	0.116*	0.019*	0.079*	0.048*	0.062*	0.063*	0.049*
	(s.e.)	(0.019)	(0.013)	(0.014)	(0.020)	(0.007)	(0.017)	(0.021)	(0.019)	(0.017)	(0.020)
Producer-to-Demand Ratio	R^2	0.708	0.630	0.779						0.237	0.769
	Demand Density	-0.313*	-0.324*	-0.294*						-0.361*	-0.242*
	(s.e.)	(0.022)	(0.018)	(0.025)						(0.037)	(0.019)
Average Plant Output	R^2	0.557	0.478	0.633		0.609				0.487	0.610
	Demand Density	0.184*	0.199*	0.163*		0.200*				0.107*	0.129*
	(s.e.)	(0.017)	(0.016)	(0.026)		(0.017)				(0.022)	(0.017)

This table shows results from robustness tests on the results presented in Table 2. See Table 5 for the key to specifications (by column). Heteroskedasticity-robust standard errors are reported. Average plant output is revenue-based in specification 5.

Table 5. Robustness Checks—Key to Specifications

[1] Benchmark (from Table 2)
[2] Minimum number of plants in market-year cell with TFP data is 2 rather than 5, N = 974
[3] Minimum number of plants in market-year cell with TFP data is 10 rather than 5, N = 343
[4] Labor productivity (logged output per worker-hour) instead of TFP, market K/L controls
[5] Revenue-based TFP measure
[6] Assumed scale elasticity of 0.9
[7] Assumed scale elasticity of 1.1
[8] Estimated TFP from production function
[9] Single cross section using average TFP and density measures, N = 312
[10] Market size control included

Table A.1. Production Function Regression Results

	First Stage		Second Stage				
N	F-Statistic for Demand Instruments	Sum of Demand Instrument Coefficients	δ_{1982}	δ_{1987}	γ_{mult}	γ_x	R^2
10,613	35.2 ($p < 0.001$)	1.025 (0.115)	-0.115 (0.010)	-0.079 (0.012)	0.070 (0.010)	0.996 (0.029)	0.902

Production Function Coefficients:
δ_{1982}—1982 year fixed effect
δ_{1982}—1987 year fixed effect
γ_{mult}—Multi-plant firm effect
γ_x—Scale elasticity

Reported standard errors are robust and clustered by CEA (local market)