Finding Mr. Schumpeter: An Empirical Study of Competition and Technology Adoption*

Jeffrey Macher† Georgetown University

Nathan H. Miller‡ Georgetown University

Matthew Osborne§ University of Toronto

July 20, 2016

Abstract

We estimate the effect of competition on the adoption of a cost-reducing technology in the cement industry, using data that span 1949-2013. The new technology, the precalciner kiln, reduces fuel usage and hence fuel costs. We find adoption is more likely if the cost savings are large, and less likely if there are many nearby competitors (accounting for the endogeneity of competition). We also find that competition damps the positive effect of cost savings. These results are consistent with a two-stage theoretical model in which firms consider adoption and then compete in quantities. We develop implications for environmental and antitrust policy.

Keywords: technology, innovation, competition, portland cement
JEL classification: L1, L5, L6

---

*We thank Philippe Aghion, Jacob Cosman, Alberto Galasso, Richard Gilbert, Arik Levinsohn, Devesh Raval, Rich Sweeney, Mihkel Tombak, Francis Vella, and seminar participants at Georgetown University, Harvard Business School, and University of Toronto for helpful comments. We have benefited from conversations with Hendrick van Oss of the USGS and a number of industry participants.

†Georgetown University, McDonough School of Business, 37th and O Streets NW, Washington DC 20057. Email: jeffrey.macher@georgetown.edu.

‡Georgetown University, McDonough School of Business, 37th and O Streets NW, Washington DC 20057. Email: nathan.miller@georgetown.edu.

§University of Toronto, Rotman School of Management, 105 St. George St., Toronto, ON, Canada M5S 3E6. Email: matthew.osborne@rotman.utoronto.ca

1 Introduction

The effect of competition on innovation has been a focus in the economics literature since at least the seminal work of Joseph Schumpeter (1934, 1942), who posited that the stability of large firms in concentrated markets gives rise to greater R&D investment. Empirical research on this subject is difficult because innovation is hard to measure and competition itself is an endogenous outcome. These concerns, however, have not deterred researchers: a vast number of articles examine descriptive regressions based on cross-industry comparisons of R&D spending and patent counts. Other contributions examine how competition and firm size affect technology adoption decisions (e.g., Oster (1982); Hannan and McDowell (1984); Rose and Joskow (1990); Schmidt-Dengler (2006)). More recently, a number of articles employ dynamic structural models to simulate the effect of competition, using theory to supplement empirical variation (e.g., Goettler and Gordon (2011); Igami (2015); Fowlie, Reguant and Ryan (2016); Igami and Uetake (2016)).[1]

We study the diffusion of precalciner kiln technology in the United States portland cement industry. Several factors make this setting amenable to empirical analysis. In data spanning 1949-2013, we have annual observations on hundreds of older kilns that are candidates to be replaced with precalciner technology. The first adoption occurs in 1974, and precalciner kilns account for the bulk of domestic capacity by the end of the sample. Most older kilns that are not replaced with precalciner technology are shut down instead. The benefits of adoption derive from reduced production costs, and can be quantified with information on fossil fuel prices and the energy efficiency of old and new technology. Changes in fossil fuel prices over the sample period create exogenous variation in the cost savings available. The high transportation costs for portland cement make competitive conditions localized and plant-specific, so the benefits of adoption experienced by plants are heterogeneous. The institutional details of the industry suggest an instrument that provides plausibly exogenous variation in local competition.

To frame the analysis, we first extend the theoretical model of Dasgupta and Stiglitz (1980) on cost-reducing investments in Cournot equilibrium. This model incorporates precalciner adoption as both non-drastic and non-divisible: non-drastic because the fuel cost savings are insufficient to price competitors out of the market; and non-divisible because the decision is fundamentally binary. Within this context, we show that competition lessens

[1] Aghion and Tirole (1994) refer to the Schumpeterian hypothesis regarding the impact of firm size and market structure on innovation as the second most tested relationship in industrial organization. Even the literature reviews are daunting (e.g., Kamien and Schwartz (1982); Baldwin and Scott (1987); Cohen and Levin (1989); Cohen (1995); Gilbert (2006); Cohen (2010)).

the adoption incentives of a myopic "focal firm" under mild conditions. We also solve a two-stage version of the model in which every firm can adopt the new technology, and show that competition limits adoption in subgame perfect equilibrium if investment costs are sufficiently large. The mechanism behind both results is deceptively simple: competition denies firms the scale required to recoup the investment costs of technology adoption. The model provides a number of other testable predictions that we take to the data.

The empirical model is based on a static perfect information game in which producers with non-precalciner kilns make a technology decision: (i) adopt precalciner technology; (ii) operate the older kiln; or (iii) shut down the older kiln without replacement. We parameterize the net benefits of adoption as depending on the number of nearby competitors, the fuel cost savings achieved with adoption, proximate construction activity, and various other controls. Theory suggests that the competition measure is positively correlated with the structural error term, which captures favorable but unobserved demand and cost conditions. This outcome introduces an endogeneity concern: to the extent that competition deters adoption, estimation risks understating the magnitude of this effect (or even suggesting that it is positive). We address endogeneity by using a 20-year competition lag as an excluded instrument, which is possible due to the length of the panel data. This instrument has power because kilns are long-lived, and is valid provided that autocorrelation in the structural errors is not too great. We estimate the model as a multinomial probit, using the two-stage conditional maximum likelihood estimator of Rivers and Vuong (1988).

The econometric results are consistent with the motivating theory. The presence of many nearby competitors diminishes the net benefits of adoption, and this effect is statistically significant, robust, and large: the mean elasticity of the adoption probability with respect to competition ranges between -1.45 and -2.16 in the baseline specifications. By contrast, the impacts of fuel cost savings and construction activity on technology adoption are positive, robust, and large: the mean elasticity of adoption with respect to fuel cost savings ranges from 0.57 and 0.82 and with respect to construction activity ranges from 1.16 and 1.54 in the baseline specifications. These magnitudes provide support for mechanism of the motivating theory: if demand and competition both increase by a similar amount then the net effect on adoption incentives is small. We also consider technology abandonment (i.e., kiln shutdowns without replacement). Our empirical results indicate that the probability of kiln shutdown increases with fuel costs and the number of nearby competitors, and decreases with proximate construction activity.

Our research has direct policy relevance in at least two arenas. First, our research is relevant to whether market-based CO_2 regulation would induce firms to adopt more efficient "green" technology. Existing empirical articles on induced innovation typically find some margin of adjustment but do not address issues of competition (e.g., Newell, Jaffe and Stavins (1999); Popp (2002); Linn (2008); Aghion et al (2012); Hanlon (2014)). Our results indicate that firms are responsive *if competition is not too great.* This interactive effect arises both in the motivating theory and in the data. Our estimates imply that a monopolist facing a one standard deviation fuel cost shock is nearly five times more likely to adopt precalciner technology in response, relative to a firm facing an average number of nearby competitors. Firms with many nearby competitors are more likely to shut down than to adopt new technology. These findings raise interesting and important questions about dynamic adjustment paths that require a more sophisticated methodology than in this paper.

Second, our research is relevant to the antitrust review of mergers. Allegations that mergers among competitors damp innovation incentives appear with some frequency in the Complaints of the DOJ and FTC (Gilbert (2006)). It is often difficult for outside economists to evaluate the merits of these allegations, because the court documents typically do not elaborate on the theoretical mechanism by which market structure affects innovation incentives. Our results suggest one specific setting in which mergers could have pro-competitive effects on innovation; for innovations that are non-drastic and non-divisible, it is possible that consolidation allows firms to achieve the scale required to profitably recoup the fixed costs of investment. Our empirical results do not inform the appropriate standard under which such an efficiency should be deemed sufficiently substantial or cognizable.

The external validity of our results is best developed via a brief literature review. It has long been understood that market power can facilitate innovation (e.g., Dasgupta and Stiglitz (1980)). Yet the opposite effect can arise if, for example, innovation cannibalizes monopoly profit (Arrow (1962)), preemptive investments deter entry (Gilbert and Newbery (1982)), or firms innovate to escape competitive pressure (Aghion et al (2005)). As many other possibilities exist, we refer interested readers to Aghion and Griffith (2005) and Gilbert (2006) for useful and complementary literature reviews. The institutional details of the market matter because there is no single theory that applies to all situations. We believe the two most important institutional details in our application are that technology adoption is both non-drastic and non-divisible. These conditions enable the existence of competitors to deprive firms of the scale necessary to recoup investment costs.

With non-drastic and non-divisible investments, it is still possible for competition to speed adoption due to preemption incentives in the medium run, even as it limits adoption in the long run.[2] We do not find support for preemption in the data – the presence of early adopters has little additional explanatory power over adoption decisions. We suspect that this result may be due to the large number of competitors that the average plant faces, which thereby distinguishes cement from others settings in which researchers find evidence of preemption that typically feature either tight oligopolies (e.g., Genesove (1999); Vogt (2000); Schmidt-Dengler (2006); Gil, Houde and Takahashi (2015)) or a monopolist that invests to deter entry (e.g., Dafny (2005); Ellison and Ellison (2011)). Our results therefore extend most naturally to technology adoption in economic settings in which Schumpeterian scale effects dominate preemption incentives.

Our research builds on the substantial literature on technology adoption. The earliest contributions study competitive environments (e.g., Griliches (1957)) and thus do not address the research questions examined here. Empirical support for the Schumpeterian prediction that firm size encourages adoption has been found in a number of settings, including ATMs and credit scoring in banking (Hannan and McDowell (1984); Akhavein, Frame and White (2005)), coal-fired steam-electric generating technologies among electric utilities (Rose and Joskow (1990)), machine tools in engineering (Karshenas and Stoneman (1993)), and MRIs in hospitals (Schmidt-Dengler (2006)). This is consistent with the mechanism of the motivating theory, in which competition disciplines firm size. There are, of course, counter-examples: perhaps most notably, the case of the basic oxygen furnace in the steel industry (e.g., Oster (1982)). Within this literature, our research is distinguished by the amount of cross-sectional and time-series variation in the cement data, as well as by the availability of an instrument that provides plausibly exogenous variation in the competitive environment.

Among the recent dynamic structural articles on competition and innovation, the most directly relevant for our research is Fowlie, Reguant and Ryan (2016). These authors simulate the effects of market-based CO_2 regulation on the portland cement industry, based on a model that allows plants to make forward-looking capacity and exit decisions. State-space payoffs are determined by Nash-Cournot competition within local markets. The simulations indicate that regulation induces exit and capacity-reductions, which comports with our econometric results that higher fuel prices increase the propensity of plants to shut down older kilns. The authors' simulation does not allow for technology adoption, however, which our results suggest is a meaningful margin of adjustment. We therefore view the two research projects as complementary. Overall, our empirical results support the usefulness of the underlying structural framework, and our theoretical extension of Dasgupta and Stiglitz (1980) helps

[2]The appendix provides a simple extension to the motivating theory that develops this point.

clarify the mechanisms through which regulation affects dynamic decisions.

The paper proceeds as follows. Section 2 develops the motivating theory, including both the focal firm model and the two-stage model. Section 3 provides institutional details on precalciner kilns and the portland cement industry. Section 4 develops the two-stage game of perfect information that we take to the data, and also discusses identification. Section 5 defines the variables used in the empirical analysis and provides summary statistics. Section 6 describes the results of the regression analysis, and Section 7 concludes.

2 Motivating Theory

The motivating theory has two parts. In the first part, a "focal firm" chooses whether to adopt a cost saving technology and then competes a la Nash Cournot with it competitors. This simple framework clarifies the mechanisms at play, and also informs our empirical specification. In the second part, all firms choose simultaneously whether to adopt the technology, and then compete a la Nash Cournot. This approach better illuminates the long run relationships. Under weak conditions, both models imply that competition reduces technology adoption. Neither model incorporates preemption incentives. In Appendix C, we provide a simple model to illustrate that preemption can cause competition to speed adoption in the "medium run" even as it limits adoption in the long run.

2.1 Focal firm model

Consider competition in a market with $j = 1, \ldots, N$ firms that have constant marginal cost functions. A single "focal firm" can purchase a technology that lowers its marginal cost by paying the capital cost k. The focal firm's marginal costs with and without technology adoption are c_i^1 and c_i^0, respectively, such that $c_i^1 = c_i^0 - \Delta c \geq 0$. The equilibrium producer surplus of the focal firm is given by the function $\pi_i(c_i, c_{-i}, \gamma)$, where c_{-i} is a vector of competitors' marginal costs and γ incorporates all other factors. The focal firm adopts the technology if the benefit exceeds the cost:

$$b_i(c_i^0, c_i^1, c_{-i}, \gamma) \equiv \pi_i(c_i^1, c_{-i}, \gamma) - \pi_i(c_i^0, c_{-i}, \gamma) > k \tag{1}$$

Assumption A1: *The equilibrium producer surplus function is differentiable.*

A1 rules out undifferentiated Nash-Bertrand competition, but holds in most other standard oligopoly models. Yet even with only this weak assumption, it is still possible to derive a result of empirical relevance. Consider a hypothetical per-unit production subsidy, s, that applies uniquely to the focal firm. The benefits of technology adoption can be reexpressed using a first order Taylor Series expansion:

$$b(\Delta c, c_i^0, c_{-i}, \gamma) \approx \left. \frac{\partial \pi_i(c_i - s, c_{-i}, \gamma)}{\partial s} \right|_{c_i = \widehat{c}_i} \Delta c \tag{2}$$

where $\widehat{c}_i = \frac{2c_i^0 - \Delta c}{2}$. To a first order approximation, competition and the other factors summarized in γ affect the benefits of technology adoption through their interaction with Δc.

Assumption A2: *Firms produce a homogeneous product and competition is Nash-Cournot. Prices are given by the inverse demand curve $P(Q) = a - Q$, for $Q = \sum_{j=1}^{N} q_j$.*

We use the Cournot model to develop theoretical results both because it yields tractable closed-form solutions with a linear demand curve and because it is a reasonable match for the portland cement industry (e.g., Ryan (2012); Fowlie, Reguant and Ryan (2016)). Equilibrium markups and quantities are given by:

$$P^*(c_j, \overline{c}, a, N) - c_j = q^*(c_j, \overline{c}, a, N) = \frac{a - c_j + N(\overline{c} - c_j)}{(N+1)} \tag{3}$$

where $\overline{c} = \frac{1}{N}\sum_j c_j$. The first equality implies that the equilibrium surplus function is given by $\pi_j^*(c_j, \overline{c}, a, N) = (q^*(c_j, \overline{c}, a, N))^2$. We restrict attention to markets in which quantities are positive for all firms. The unit slope normalization is without loss of generality, as all results derived below extend easily to demand curve rotations. We refer readers to Tombak (2002) for a derivation of these equilibrium objects, and to Shapiro (1989) for a more general discussion of the Nash-Cournot model, including conditions for the existence and uniqueness of equilibrium with nonlinear demand.

The benefit that focal firm i receives from technology adoption is given by:

$$b_i(\Delta c, c_i^0, \overline{\overline{c}}, a, N) = \frac{2N}{(N+1)} q^*(\widehat{c}_i, \overline{\overline{c}}, a, N)\Delta c \tag{4}$$

where $\overline{\overline{c}} = \frac{1}{N}(\sum_{j \neq i} c_j^0 + \widehat{c}_i)$ is the average marginal cost evaluated at the midpoint between the focal firm's new and old cost. The equation is derived directly from the equilibrium markups and quantities, and found in Appendix B. Alternatively, because the producer surplus of the focal firm is quadratic in its costs, the Taylor series approximation in equation (2) holds with equality and equation (4) can be obtained by differentiating the producer surplus function with respect to the magnitude of a per-unit production subsidy.

It is straightforward to show that the benefit of adoption increases with consumer willingness-to-pay (i.e., a) and with the cost savings (i.e., Δc): the former effect arises because a increases equilibrium quantity; the latter effect is seen by conceptualizing an increase in Δc due to symmetric changes in c^1 and c^0, so that $\widehat{c}_i$ and $\bar{\bar{c}}$ are unchanged, in which case the relevant derivative is unambiguously positive.

Result 1: *Under A1 and A2, the benefits of technology adoption:*

(i) increase with demand: $\frac{\partial b_i(\Delta c, c_i^0, \bar{\bar{c}}, a, N)}{\partial a} > 0$

(ii) increase with the cost savings: $\frac{\partial b_i(\Delta c, c_i^0, \bar{\bar{c}}, a, N)}{\partial \Delta c} > 0$

To assess how the level of competition (i.e., N) affects the benefits of adoption, it is necessary to specify the marginal costs of the firms being added or removed from the market. Define $\bar{\bar{c}} = \frac{1}{N}(\sum_{j \neq i} c_j^0 + \widehat{c}_i)$ as the market's average marginal cost evaluated at the midpoint between the focal firm's new and old cost, and consider the thought experiment of adding a new firm with a marginal cost equal to $\bar{\bar{c}}$. This change decreases the benefit that the focal firm receives from adoption if the following condition holds:

$$a - \widehat{c}_i > (\bar{\bar{c}} - \widehat{c}_i)\left[\frac{N^2(N+2)^2 - (N+1)^4}{(N+1)^3 - N(N+2)^2}\right] \tag{5}$$

The term in brackets equals 3.4 if $N = 2$, and converges quickly to two as N grows large. Appendix B.1 provides a proof. Condition (5) is likely to hold in the empirical application because recent empirical research finds that average markups, which provide a lower bound to $(a - \widehat{c}_i)/\widehat{c}_i$, are 130 percent in the cement industry (Ganapati, Shapiro and Walker (2016)).[3] Under this condition, the interactive effect applies such that competition reduces the (positive) effect that cost savings have on adoption.

Result 2: *Under A1, A2, and condition* (5)*, the benefits of technology adoption decrease in the number of competitors:* $\frac{\Delta b_i(\Delta c, c^0, a, L, N)}{\Delta N} < 0$*. Increasing the number of competitors reduces the (positive) derivative of adoption benefits with respect to* Δc*.*

Finally, to build intuition on why the benefit of technology adoption decreases with N, reconsider the approximation in equation (2). Under Nash-Cournot competition but

[3]The authors infer markups based on data from the Census Bureau's Census of Manufactures. See also Section 3, where we show that prices well exceed fuel costs.

relaxing the linearity demand assumption, the derivative of the producer surplus function can be decomposed as follows:

$$\frac{\partial \pi_i(c_i - s, c_{-i}, \gamma)}{\partial s} = \underbrace{q_i^*(c_i, c_{-i}, \gamma)}_{\text{Cost Savings}} \underbrace{- q_i^*(c_i, c_{-i}, \gamma) \sum_{k \neq i} \frac{\partial P^*(Q)}{\partial q_k} \frac{\partial q_k^*(c_k, c_{-k}, \gamma)}{\partial q_i} \frac{\partial q_i^*(c_i, c_{-i}, \gamma)}{\partial s}}_{\text{Strategic Effect}}$$

The first term (a cost savings effect) represents how lower marginal cost increases the variable profit for each unit produced (Arrow (1962)). This effect provides the main mechanism through which increasing N reduces the adoption incentive: more competitors means lower equilibrium output, which reduces the available cost savings. The second term (a strategic effect) represents how the focal firm – given its lower marginal costs – induces competitors to produce less, which subsequently raises price. With linear demand, the strategic effect simplifies to $q_i^*(c_i, c_{-i}, a, N)\left(\frac{N-1}{N+1}\right)$. Increasing N reduces the cost saving effect but can amplify the strategic effect: condition (5) determines which change dominates.

2.2 Two-stage model

We now extend the motivating theory to a setting in which all firms have the opportunity to adopt a cost reducing technology. Specifically, we examine a game with two stages. In the first stage, each of N firms can adopt a technology by paying the capital cost k. In the second stage, all firms compete according to undifferentiated-products Nash-Cournot. Marginal costs with and without the technology are c_1 and c_0, respectively, and again $c^1 = c^0 - \Delta c \geq 0$. Prices are given by the inverse demand curve $P(Q) = a - Q$, for $Q = \sum_{j=1}^{N} q_j$.

We characterize the number of firms that adopt the technology in subgame perfect equilibrium (SPE). We first summarize second stage payoffs. The stage-game equilibrium markups and quantities of a non-adopter are:

$$P^*(L) - c^0 = q^*(c^0, L) = \frac{a - L\Delta c - c_0}{N+1} \tag{6}$$

where we use L to denote the number of adopters. This result is obtained by manipulating equation (3). The solutions for adopters can be shown to be given by $q^*(c^1, L) = q^*(c^0, L) + \Delta c$. Adopters thus have greater quantity and larger markups, and both effects increase with the cost savings available with the new technology.

We next turn to the first stage. The SPE is characterized by some number of adopters, $L^* \leq N$, such that (i) all adopters prefer adoption to non-adoption, and (ii) all non-adopters prefer non-adoption to adoption. Recalling that producer surplus in the second stage equals the square of equilibrium quantity, these stability conditions generate two inequalities that can be manipulated to solve the game:

$$k \leq q^*(c^1, L = L^*)^2 - q^*(c^0, L = L^* - 1)^2 \tag{7}$$

and

$$k > q^*(c^1 | L = L^* + 1)^2 - q^*(c^0 | L = L^*)^2 \tag{8}$$

Plugging in for the equilibrium quantities, the following expression obtains with some algebra that we defer to Appendix B.2:

$$L^* \leq N + \frac{a - c^0}{\Delta c} - \frac{1}{2} \frac{k}{(\Delta c)^2} \frac{(N+1)^2}{N} < L^* + 1 \tag{9}$$

By inspection, the SPE number of adopters can increase or decrease in N. This ambiguity arises because more firms simultaneously damps adoption incentives but increases the pool of possible adopters: the net effect depends on the parameter values. If capital costs are high enough relative to the cost savings (specifically, if $k > 2(\Delta c)^2$), however, then the SPE number of adopters approaches zero as the number of firms grows large. Also relevant here is the fraction of firms that adopt in the SPE, which can be obtained by dividing equation (9) by N:

$$\frac{L^*}{N} \leq 1 + \frac{1}{N} \frac{a - c^0}{\Delta c} - \frac{1}{2} \frac{k}{(\Delta c)^2} \frac{(N+1)^2}{N^2} < \frac{L^* + 1}{N} \tag{10}$$

This fraction is less than one for sufficiently high N, and approaches zero as N grows large under the same condition that $k > 2(\Delta c)^2$. Figure 1 plots the number and fraction of adopters under one such parameterization ($a = 20, k = 4, \Delta c = 1$). For $N < 7$, the number of adopters grows with the number of firms because all firms find it profitable to adopt. The number of adopters then shrinks for $N > 7$, and equals zero for $N >= 14$.

Result 3: *Under the condition $k > 2(\Delta c)^2$, there exists some N_1 and N_2 ($N_1 < N_2$) such that (1) if $N > N_1$ then the fraction of firms that adopt the technology in SPE decreases with N and (2) if $N > N_2$ then no firms adopt.*

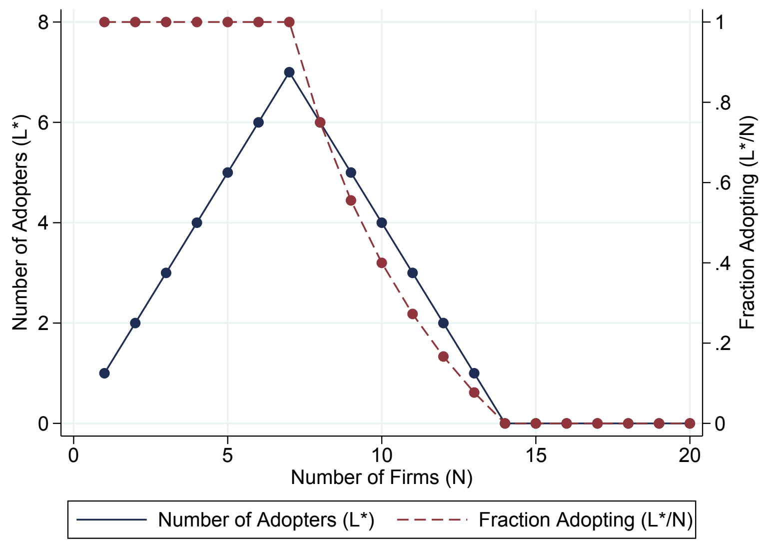

Figure 1: Technology Adoption in the Two-Stage Model

3 Empirical Setting

3.1 The portland cement industry

We examine the adoption of precalciner technology in the portland cement industry over 1973-2013. Portland cement is a finely ground dust that forms concrete when mixed with water and coarse aggregates such as sand and stone. Concrete, in turn, is an essential input to many construction and transportation projects. The production of cement involves feeding limestone and other raw materials into rotary kilns that reach peak temperatures of 1400-1450° Celsius. Because the associated fuel costs account for a sizable portion of revenues, the fuel efficiency of the kiln technology employed is an important determinant of a plant's overall profitability.

Plants equipped with precalciner technology preheat the raw materials using the exhaust gases of the kiln and heat supplied by a supplementary combustion chamber. This approach reduces the energy requirements of production by 25-35 percent relative to older wet and long dry kilns. Installation of precalciner technology requires a retrofit of the entire kiln system. Because the process allows one of the main chemical reactions to occur before raw materials enter the kiln, the requisite kiln length is greatly reduced. The cost of installation is large, however, and as a result adoption of the technology has been gradual since its development in the late 1960s and early 1970s.

Table 1 tracks precalciner kiln adoption over time. In 1973, nearly all plants used

Table 1: The Portland Cement Industry over 1973-2013

Year	Wet Kilns	Long Dry Kilns	Dry with Preheater	Dry with Precalciner	Total Kilns	Total Plants	Total Capacity
1973	249	157	23	0	429	159	76.67
1978	201	111	42	2	356	151	79.85
1983	121	90	36	24	271	132	79.79
1988	96	70	35	26	227	116	75.47
1993	72	65	38	27	202	107	74.50
1998	67	63	34	31	195	106	76.79
2003	53	49	38	45	185	106	90.88
2008	45	31	32	56	164	103	96.00
2013	19	26	29	66	140	95	98.45

Notes: The table shows data at five-year snapshots spanning 1973-2013. Kiln counts are provided separately for each of the four production technologies: wet kiln, long dry kilns, dry kilns with preheaters, and dry kilns with precalciners. Total capacity is in millions of metric tonnes. The data are for the contiguous U.S. and are obtained from the *PCA Plant Information Survey*.

inefficient wet and long dry kilns. A small number used preheater technology, which recycles exhaust gases without a supplementary combustion chamber, but none used precalciners. Over the ensuing four decades, the number of wet kilns decreased from 249 to 19, the number of long dry kilns decreased from 157 to 26, and the number of precalciner kilns increased substantially. In the final year of data, 66 of the 140 kilns in operation use precalciners and account for 74 percent of industry capacity. Indeed, the higher capacity of precalciner kilns explains why industry capacity increased as the total number of plants and kilns decreased.

Table 2 provides the average fuel costs among kilns in each technology class, again at five-year intervals over the sample period. These costs are obtained based on kiln efficiency and the price/mBtu of the primary fossil fuel used. The changes within kiln technology classes over time are driven primarily by exogenous fluctuations in natural gas and coal prices, which provides a key source of variation that we exploit in the estimation. This feature of the data can be interpreted further as providing a natural experiment as to how firms would respond to the market-based regulation of CO_2 emissions (which would change the implicit price of fossil fuels).

Table 2 also provides the fuel costs of the "frontier technology," which we define as a precalciner kiln that burns the most affordable fuel. The difference between a kiln's fuel cost and that of the frontier technology – a measure of the fuel cost savings available from precalciner adoption – is an empirical analog to the Δc term in the motivating theory. Fuel

Table 2: Fuel Costs per Metric Tonne of Cement

Year	Wet Kilns	Long Dry Kilns	Dry with Preheater	Dry with Precalciner	Frontier Technology	Average Price
1973	18.99	16.41	13.30	·	10.33	85.59
1978	36.42	31.13	24.56	23.35	22.41	110.25
1983	28.84	23.63	18.06	16.78	16.78	94.41
1988	19.81	15.91	13.28	12.41	11.75	79.78
1993	15.35	12.66	9.86	9.77	9.44	77.97
1998	13.50	11.24	8.75	8.39	8.39	98.13
2003	12.94	11.26	8.76	8.40	8.40	87.53
2008	22.81	19.85	15.45	14.81	14.81	105.55
2013	25.70	22.36	17.40	16.83	16.68	89.93

Notes: The table provides average fuel costs by kiln technology, the hypothetical fuel costs of a kiln with "frontier technology" defined as a precalciner kiln that burns the most affordable fuel, and the national average price of portland cement. Data are shown at five-year snapshots spanning 1973-2013. Fuel costs are based on authors' calculations as detailed in Appendix A. Prices are obtained from the USGS *Minerals Yearbook*. All statistics are in real 2010 dollars per metric tonne of cement output.

cost savings tend to be large when fossil fuel prices (and thus fuel costs) are high.[4] The final column of the table provides the national average price of portland cement: depending on the year and kiln technology, fuel costs account for between 8 to 33 percent of revenues. Two recent papers estimate that pass-through of fuel costs to price in the industry exceeds unity (Miller, Osborne and Sheu (2015); Ganapati, Shapiro and Walker (2016)).

Cement is typically transported by truck to ready-mix concrete plants and large construction sites, and these associated costs generally account for a sizable portion of purchasers' total expenditures. The academic literature often models the industry as a number of distinct local markets (e.g., Ryan (2012); Fowlie, Reguant and Ryan (2016)). Figure 2 provides a map of the cement plants in operation as of 2010. Some geographic areas (e.g., southern California) have many plants, while others areas (e.g., South Dakota) have only a single nearby plant. These differences provide cross-sectional variation in competitive conditions.

Cement is used in construction projects and demand is highly procyclical as a result. Figure 3 graphs total production and consumption in the United States over 1973-2013. When macroeconomic conditions are favorable, consumption tends to outstrip production due to domestic capacity constraints; imports make up the differential. The technology

[4]There is a well known analogy in the automobile industry: the driving cost of vehicles with low miles-per-gallon (MPG) is more sensitive to the gasoline price than that of high MPG vehicles, and automobile prices adjust accordingly (e.g., Busse, Knittel and Zettelmeyer (2013); Langer and Miller (2013)).

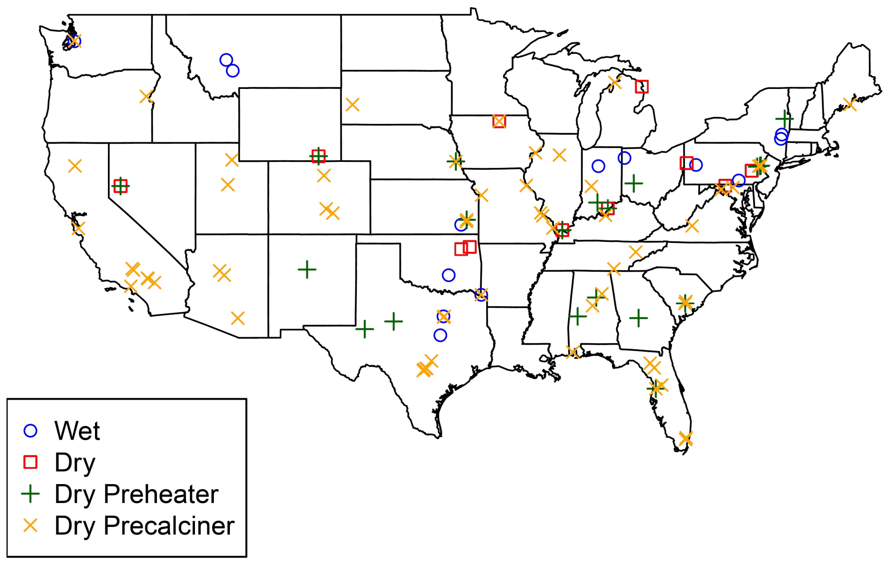

Figure 2: Portland Cement Plants in the Contiguous United States, in 2010

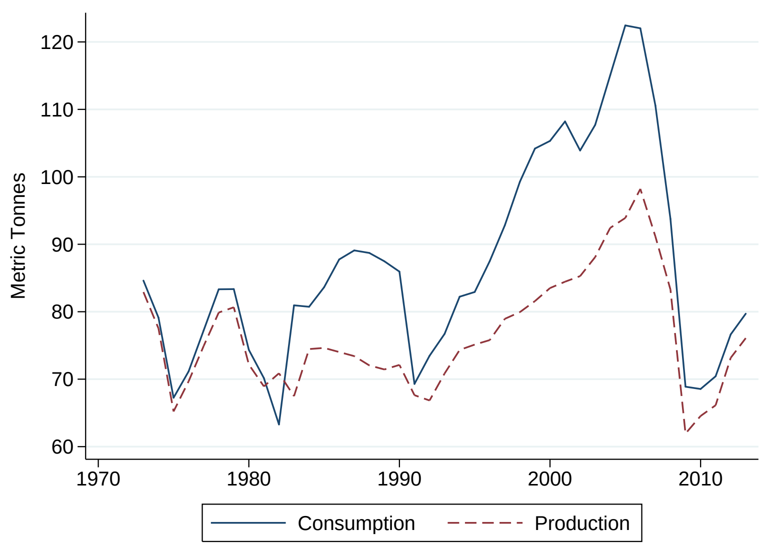

Figure 3: Consumption and Production in the United States, 1973-2013

by which cement can be shipped via transoceanic freighter at low cost and imported was developed in the late 1970s, which explains the tight connection between consumption and production in the earliest years of the sample. U.S. cement exports are negligible. Finally, cement cannot be stored for any meaningful period of time, because the product gradually absorbs moisture in the air which eventually renders it unusable.

3.2 Data sources

We draw on several data sources to construct a panel of kiln-year observations that span the contiguous United States over 1973-2013. This sample period is determined by the Portland Cement Association's (PCA) *Plant Information Survey* (PIS), published annually over 1973-2003, semi-annually over 2004-2010, and then again in 2013. The PIS provides the location, owner, and primary fuel of each cement plant in the U.S. and Canada, as well as the age, capacity and technology class of each kiln. We impute values in missing years by using data from preceding and following years, as well as by using information in the *Minerals Yearbook* of the United States Geological Survey (USGS), which provides an annual cement plant census. We combine the PIS kiln data with supplementary data that contain kiln locations over 1949-1973. These data were constructed by backcasting the 1973 PIS using information culled from the trade publication *Pit and Quarry*, occasionally printed *Pit and Quarry* maps of the industry, and the *American Cement Directory*. We refer readers to Chicu (2012) for details.[5] The supplementary dataset is useful because it allows us to construct lagged competition measures without discarding the earlier years of the PIS sample.

We calculate the fuel costs of production based on kiln efficiency and fossil fuel prices, using the PCA's *U.S. and Canadian Portland Cement Labor-Energy Input Survey* to measure production energy requirements. This survey is published intermittently, and we use the 1974-1979, 1990, 2000, and 2010 versions. We obtain the average prices of coal, natural gas, and distillate fuel oil for the industrial sector from the State Energy Database System (SEDS) of the Energy Information Agency (EIA). We use fossil fuel prices at the national level because they are more predictive of cement prices (Miller, Osborne and Sheu (2015)), probably due to the measurement error associated with imputing withheld state-level data. We obtain retail gasoline prices from the EIA's *Monthly Energy Review*.[6] We use county-level data on construction employment and building permits from the Census Bureau to account for demand-side fluctuations.[7] Construction employment is part of the County Business Patterns data. We use NAICS Code 23 and (for earlier years) SIC Code 15. The

[5] We thank Mark Chicu for making these data available.

[6] The gasoline prices include federal and sales taxes, and are for regular leaded gasoline until 1990 and regular unleaded gasoline thereafter. See `http://energy.gov/eere/vehicles/fact-915-march-7-2016-average-historical-annual-gasoline-pump-price-1929-2015`, last accessed April 25, 2016.

[7] For both the construction employment and building permits, it is necessary to impute a small number of missing values. We calculate the average percentage difference between the observed data of each county and the corresponding state data, and use that together with the state data to fill in the missing values.

data for 1986-2010 are available online.[8] The data for 1973-1985 are obtained from the University of Michigan Data Warehouse. The building permits data are maintained online by the U.S. Department of Housing and Urban Development.[9] Finally, data on cement prices, consumption, and production reported in the previous subsection are obtained from the USGS *Minerals Yearbook*. USGS does not provide firm-level or plant-level data.

4 Empirical Model

We estimate a two-stage game of perfect information. In the first stage, producers with non-precalciner kilns make a technology decision to (i) adopt precalciner technology; (ii) operate the older kiln; or (iii) shut down the older kiln without replacement. In the second stage, all producers compete taking the outcomes of the first stage as given. The payoff functions are parameterized based on the motivating theory. We conceptualize producers as playing this two-stage game each year, which exploits the panel data setting. This framing is typical of the static perfect information games estimated in the literature (e.g., Bresnahan and Reiss (1991); Berry (1992); Gowrisankaran and Stavins (2004); Toivanen and Waterson (2005); Ciliberto and Tamer (2009); Perez-Saiz (2015)).

4.1 Payoffs

We assume that if producer i adopts precalciner technology in year t then its profit is $\pi(c^1_{it}, N_{it}, a_{it}, w_{it}; \theta) - k_t + u^A_{it}$, where c^1_{it} is the fuel costs with precalciner technology, N_{it} is a continuous measure of the competition faced by the producer, a_{it} is a measure of demand, w_{it} is a vector of controls, k_t is the capital cost of adoption, and u^A_{it} is a stochastic term. If the producer instead maintains the old kiln, its profit is $\pi(c^0_{it}, N_{it}, a_{it}, w_{it}; \theta) + u^0_{it}$. The profit of shutting down the old kiln is u^S_{it}, which can be interpreted as a stochastic draw on exit costs. Producers make their technology decision to maximize Π^*_{it}, which is given by:

$$\Pi^*_{it} = \begin{cases} b(\Delta c_{it}, c^0_{it}, N_{it}, a_{it}, w_{it}; \theta) - k_t + u^A_{it} & \text{if adopt} \\ u^0_{it} & \text{if maintain} \\ -\pi(c^0_{it}, N_{it}, a_{it}, w_{it}; \theta) + u^S_{it} & \text{if shut down} \end{cases} \tag{11}$$

where $b(\Delta c_{it}, c^0_{it}, N_{it}, a_{it}, w_{it}; \theta) = \pi(c^1_{it}, N_{it}, a_{it}, w_{it}; \theta) - \pi(c^0_{it}, N_{it}, a_{it}, w_{it}; \theta)$. Formulating the choice in this manner aligns the empirical model with the motivating theory. The technology

[8] See `http://www.census.gov/econ/cbp/download/`, last accessed April 16, 2014.

[9] See `http://socds.huduser.org/permits/`, last accessed April 16, 2014.

decisions are observable even though Π_{it}^* is not, and the dependent variable is Π_{it}, which equals 0, 1, and 2, for adopt, maintain, and shut down, respectively.

We parameterize the profit functions using linear approximations. The profit of adopting precalciner technology, relative to maintaining the old kiln, is given by

$$b_{it} - k_{it} + u_{it}^A \equiv y_{it}^A = \beta_1^A \Delta c_{it} + \beta_2^A N_{it} + \beta_3^A a_{it} + \beta_4^A c_{it}^0 + w_{it}' \alpha^A + \phi_t^A + u_{it}^A \tag{12}$$

where ϕ_t^A allows for changes over time. The empirical variation is insufficient to separately identify β_1^A and β_4^A, so we impose a normalization ($\beta_4^A = 0$) and focus on arguably the more important effect. The profit of technology abandonment (i.e., kiln shutdown) also has a linear specification:

$$-\pi(c_{it}^0, N_{it}, a_{it}, w_{it}; \theta) + u_{it}^S \equiv y_{it}^S = \beta_1^S c_{it}^0 + \beta_2^S N_{it} + \beta_3^S a_{it} + w_{it}' \alpha^S + \phi_t^S + u_{it}^S \tag{13}$$

The model thus indicates that Δc_{it} enters the adoption equation but not the shutdown equation. We specify ϕ_t^A and ϕ_t^S alternately using linear time trends, a flexible polynomial in time, and year fixed effects. The control variables include kiln age and kiln capacity.

4.2 Timing and equilibrium

The uniqueness of equilibrium in static games of perfect information is not guaranteed. For example, it is possible in our application that the market can support a certain number of firms who adopt precalciner technology, but that the identities of the adopters in equilibrium is indeterminate. This potential can be problematic for identification and estimation because there may not be a one-to-one mapping between the model primitives and outcomes (e.g., Bajari, Hong and Ryan (2010)). To ensure uniqueness, we assume that decisions are made sequentially in the order of kiln age, with older kilns moving first. This approach is reasonable because the econometric evidence supports that older kilns are more likely to be replaced with precalciner technology. Similar timing assumptions are employed in the literature (e.g., Toivanen and Waterson (2005)).

4.3 Estimation

We estimate the model using multinomial probit regression. Because the measure of competition, N_{it}, is potentially correlated with the structural error terms, we employ the two stage conditional maximum likelihood estimator developed by Rivers and Vuong (1988). This estimator requires a reduced-form equation to govern the evolution of the endogenous variable. We assume that N_{it} evolves according to the following equation:

$$N_{it} = z_{it}\gamma_1 + \Delta c_{it}\gamma_2 + c^0_{it}\gamma_3 + a_{it}\gamma_4 + w'_{it}\gamma_5 + \phi^N_t + v_{it} \tag{14}$$

where z_{it} is an instrument that is excluded from the structural equations, ϕ^N_t is specified the same way as ϕ^A_t and ϕ^S_t, and v_{it} is a reduced-form error term. For notational convenience, we collect the exogenous variables (z_{it}, Δc_{it}, c^0_{it}, w_{it}) in the vector X_{it}. We assume that $(X_{it}, u^A_{it}, u^S_{it}, v_{it})$ is i.i.d. Further, let $(u^A_{it}, u^S_{it}, v_{it})$ have a mean-zero joint normal distribution, conditional on X_{it}, with the finite positive definite covariance matrix:

$$\Omega \equiv \begin{bmatrix} \sigma^A_{uu} & \sigma^{AS}_{uu} & \sigma^A_{vu} \\ \sigma^{AS}_{uu} & \sigma^S_{uu} & \sigma^S_{vu} \\ \sigma^A_{vu} & \sigma^S_{vu} & \sigma_{vv} \end{bmatrix} \tag{15}$$

Endogeneity is present if the reduced-form error term is correlated with the structural error terms (specifically, if $\sigma^A_{vu} \neq 0$ or $\sigma^S_{vu} \neq 0$). Using the joint normality assumption, however, equations (12) and (13) can be rewritten as:

$$y^A_{it} = \beta^A_1 \Delta c_{it} + \beta^A_2 N_{it} + \beta^A_3 a_{it} + w'_{it}\alpha^A + \phi^A_t + v_{it}\lambda^A + \eta^A_{it} \tag{16}$$

$$y^S_{it} = \beta^S_1 c^0_{it} + \beta^S_2 N_{it} + \beta^S_3 a_{it} + w'_{it}\alpha^S + \phi^S_t + v_{it}\lambda^S + \eta^S_{it} \tag{17}$$

where $\lambda^k = \sigma^k_{vu}/\sigma_{vv}$ and $\eta^k_{it} = u^k_{it} - v_{it}\lambda^k$ for $k \in \{A, S\}$. If a suitable control function is used as a proxy for the reduced-form error, v_{it}, then the measure of competition is orthogonal to the remaining error terms. This approach is motivated by the method of Rivers and Vuong (1988) as follows:

1. Use OLS to regress N_{it} on z_{it}, Δc_{it}, c^0_{it}, a_{it}, w_{it}, and ϕ^N_t. This obtains an estimate of the reduced-form error term:

$$\widehat{v}_{it} = N_{it} - z_{it}\widehat{\gamma}_1 - \Delta c_{it}\widehat{\gamma}_2 - c^0_{it}\widehat{\gamma}_3 - a_{it}\widehat{\gamma}_4 - w'_{it}\widehat{\gamma}_5 - \widehat{\phi}^N_t$$

2. Estimate the multinomial probit model of equations (16) and (17) with maximum likelihood, using $\widehat{v}_{it}$ as a control function to account for endogeneity. This approach obtains consistent estimates of the structural parameters (β^k, α^k) for $k \in \{A, S\}$. Differences between v_{it} and $\widehat{v}_{it}$ are normally distributed and consistent with the distributional assumptions of the multinomial probit model. The estimates of λ^A and λ^S can be used

to test for the exogeneity of the competition measure.

The second-stage standard errors can be adjusted to account for the presence of the estimation of the control function using a multi-step procedure based on the minimum distance estimator of Amemiya (1978) and the steps described in Newey (1987). This adjustment has virtually no effect in our application, however, so we report the simpler unadjusted standard errors. It also is possible to cluster standard errors at the kiln-level as an *ad hoc* correction for autocorrelation, but this too has little effect on the magnitudes of the standard errors.

4.4 IV strategy

The structural error terms in equation (12) summarize the net effect of unobserved demand and cost factors. If these unobservables correlate with the number of competitors, then estimates obtained with basic probit regression are biased. To push further, the motivating theory indicates that technology adoption is more likely under favorable conditions (e.g., high demand, low marginal costs, or low capital costs), but more favorable conditions support more competitors in standard models so any correlation between u_{it} and v_{it} is likely positive. This result allows the bias to be signed: a basic probit estimator is likely to understate the extent that competition deters technology adoption.

Finding a good set of instruments to correct this endogeneity bias is not straightfoward because the analysis entails estimating profit functions. The setting differs from more standard applications in industrial organization that require the estimation of either demand or supply, and for which cost or demand shocks respectively are valid instruments. Both demand and cost variables enter in the profit function in our application so neither can be used as instruments. Our instrument is instead a lagged version of the competition measure, and if the structural error terms do not exhibit autocorrelation then this instrument is valid. This instrument has power in the first stage regression, furthermore, because kilns tend to operate for many years (e.g., the average age of a kiln in its final year is 40 years). A potential problem with this approach is that unobserved profit shocks may be persistent over time. Because we use long lags of the competition measure (up to 20 years), it is plausible that any such inter-temporal correlations will have died out by that time.

Other sources of endogeneity seem unlikely. Technology decisions within a market are not likely to drive cement demand, as it represents a small fraction of total construction costs; hence, exogeneity of the demand controls is likely to be reasonable. Endogeneity in fossil fuel prices could arise if increases in fuel demand from cement plants led to price increases in the fuel market. Any such feedback should be small because cement accounts for a fraction of the fossil fuels used in the United States. Consistent with this, bituminous coal prices do not exhibit the same pro-cyclical variation as cement demand.

5 Variables and summary statistics

5.1 Variables

We calculate the fuel costs of each kiln based on its energy requirements and the price of the primary fuel:

$$\text{Fuel Cost}_{jt} = \text{Primary Fuel Price}_{jt} \times \text{Energy Requirements}_{jt}$$

where the fuel price is in dollars per mBtu and the energy requirements are in mBtu per metric tonne of clinker. We obtain the energy requirements from the PCA labor-energy input surveys. Details on this calculation are provided in Appendix A. The cost savings that would be realized by adopting precalciner technology are the difference between the fuel costs of the kiln and those of the technology frontier, which we define based on the energy requirements of a precalciner kiln fired with the most affordable fuel. This difference provides the empirical proxy for the Δc term that appears in the motivating theory.

We measure competition based on plant locations and gasoline prices (which scale transportation costs). We define a distance metric as the multiplicative product of miles and a gasoline price index that equals one in the year 2000. For each kiln, we then define the variable *Competitors* as the number of competing plants within a radius of 400. The radius is motivated by prior findings that 80-90 percent of portland cement is trucked less than 200 miles (Census Bureau (1977); Miller and Osborne (2014)), so that plants separated by more than 400 miles are unlikely to compete for customers.[10] We exclude plants owned by the same firm from the competition measure (few such plants exist within the specified radius). We use instruments based on the locations of plants 20 years prior to the observation in question. Because gasoline prices are plausibly exogenous, we use the same distance radii to calculate the competition and lagged competition measures. To illustrate, consider a kiln observation in the year 2000, when the gasoline index equals one. The instruments

[10] Our treatment of distance reflects the predominant role of trucking in distribution. A fraction of cement is shipped to terminals by train (6 percent in 2010) or barge (11 percent in 2010), and only then is trucked to customers. Some plants may therefore be closer than our metric indicates if, for example, both are located on the same river system. Straight-line miles are highly correlated with both driving miles and driving time and, consistent with this, previously published empirical results on the industry are not sensitive to which of these measures is employed (e.g., Miller and Osborne (2014)).

are constructed based on the plants in 1980 within 400 miles of the kiln's location, even though the 1980 gasoline index differs than one. We are able to calculate lagged competition measures in this manner even for kilns that are not present in the data 20 years prior.

Finally, we control for kiln age, kiln capacity, and demand conditions. The first two controls are straight-forward and obtained from the PIS kiln data. The third control uses county-level data on building permits and construction employment, which explains nearly 90 percent of the variation in USGS-reported state-level consumption. To obtain a single control variable, we first create a county-specific demand variable as a linear combination of building permits and construction employment. The specific formula, which we estimate based on the state-level regressions, is $DEMAND = 0.0154 \times PER + 0.0122 \times EMP$, where PER and EMP are building permits and construction employment, respectively. We then sum the demand among counties within the distance radii from each kiln. We refer to this control variable as *Construction*. The motivating theory indicates that technology adoption should be more prevalent when demand conditions are favorable.

5.2 Summary statistics

Table 3 provides information on the composition of the sample. The data include observations on 460 distinct non-precalciner kilns: 144 are replaced with precalciner technology, 244 are closed without replacement, and 72 survive to the end of the sample. A kiln that is replaced or shut down exits the sample but continues to affect the *Competition* variable for the kilns that remain in the sample. The median kiln is observed for 12 years. At the median, kilns that are replaced with precalciner technology are observed for eight years, kilns that are shut down are observed for ten years, and kilns that survive to the end of the sample are observed for 41 years. There is some variation in the number of observations for surviving kilns due to (infrequent) greenfield entry. In total, there are 8,192 kiln-year observations in the regression sample.

Table 4 provides summary statistics for the dependent variables (indicators for adoption and shutdown) and the explanatory variables. Precalciner adoption and kiln shutdown are rare events: the mean of the indicators imply an empirical probability of 1.8% and 3.0%, respectively. These low probabilities are due to kiln longevity, which also is the source of the instrument's power. By construction, there is a high degree of collinearity (0.89) between *Fuel Cost* and *Cost Savings*: sufficiently large such that we are able to identify only the net effect in our regressions and motivates the exclusion of *Fuel Cost* from the adoption equation. There is no apparent bivariate relationship between adoption/shutdown and *Competitors*,

Table 3: Number of Observations per Kiln

	Count	Mean Obs.	Order Statistics: Observations				
			10%	25%	50%	75%	90%
All Kilns	460	17.81	2	6	12	34	41
Replaced Kilns	144	15.39	2	5	8	28	34
Shut Down Kilns	244	12.82	2	4	10	16	36
Surviving Kilns	72	37.57	37	41	41	41	41

Notes: The table provides the count of unique non-precalciner kilns in the 1973-2013 data, both together and separately for (i) kilns replaced with a precalciner kiln, (ii) kilns closed without replacement, and (iii) kilns in operation as the end of sample period. The table also summarizes the distribution of (annual) observations per kiln.

because the competition measure is correlated with *Construction* and theory indicates these variables should have opposite effects. This collinearity is nevertheless not strong enough to raise concerns in the multivariate regression analysis.[11] The summary statistics suggest that older kilns and smaller kilns are more likely to results in adoption/shutdown, and those relationship survive in the multivariate analysis.

Finally, we provide additional information on the empirical distribution of the competition measure because it plays an important rule in the empirical analysis. Figure 4 provides separate histograms of *Competitors* for each decade in the data. Cross-sectional variation is substantial due to the dispersion in plant locations mapped earlier, while inter-temporal variation arises due to gasoline price fluctuations and the net decrease in the number of plants over the sample period.

6 Results

6.1 Main results

Table 5 presents the baseline multinomial probit results. Panel A addresses the likelihood of precalciner kiln adoption, and Panel B addresses the likelihood of shutdown (both panels are relative to the alternative of maintaining the older kiln). The columns control for changes

[11] One way to assess whether collinearity is potentially problematic is to calculate the variance inflation factors (VIFs) of the regressors. This is done by regressing each regressor k on the other regressors, and calculating $VIF(k) = \frac{1}{1-R^2}$. The VIF for *Fuel Costs* ranges from 6.25 to more than 100 if *Cost Savings* also is a regressor, depending on how we control for time effects. A rule of thumb is that collinearity is a threat to asymptotic consistency if the VIF exceeds ten (Mela and Kopalle (2002)). No other regressor comes close to this threshold. As one example, the VIF of *Competitors* ranges from 2.63 to 3.77 even though its correlation coefficient with *Construction* is a reasonably high 0.71.

Table 4: Summary Statistics

Regressor		Mean	St. Dev	Correlation Coefficients (1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	Adoption	0.018	0.131							
(2)	Shutdown	0.030	0.170	-0.02						
(3)	Fuel Cost	22.15	9.63	0.07	0.06					
(4)	Cost Savings	7.78	6.62	0.08	0.05	0.89				
(5)	Competitors	20.56	12.34	-0.00	-0.00	0.03	0.13			
(6)	Construction	12.85	8.85	0.00	-0.04	-0.37	-0.20	0.71		
(7)	Kiln Age	30.87	16.12	0.08	0.09	-0.17	-0.09	-0.13	0.02	
(8)	Kiln Capacity	0.26	0.18	-0.05	-0.10	-0.20	-0.20	-0.01	0.13	-0.39

Notes: The table provides means, standard deviations, and correlation coefficients for the dependent variables (indicators for adoption and shutdown) and the main regressors. The regression sample is comprised of 8,192 kiln-year observations over the period 1973-2013. Capacity is in millions of metric tonnes per year.

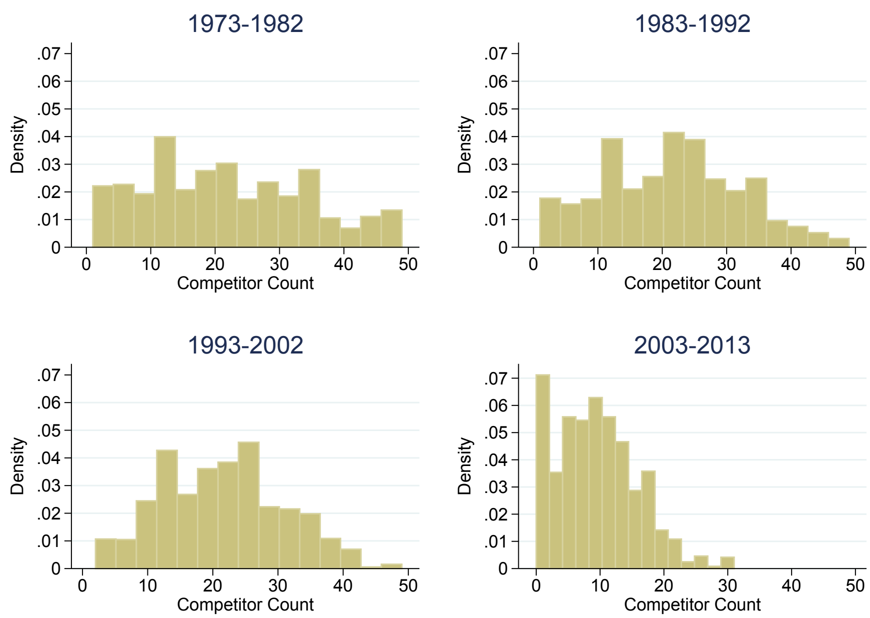

Figure 4: Count of Competitors within a Distance of 400 by Decade

over time in different ways: Column (i) relies exclusively on the regressors; column (ii) adds a linear time trend ($t = 0, 1, \ldots, 40$); column (iii) uses a fifth order polynomial in time; and column (iv) incorporates year fixed effects. The multinomial probit does not converge if year fixed effects are used, so we report the results of two binomial probit regressions.

Table 5: Probit Regression Results

	Panel A: Adopt vs. Maintain				Panel B: Shut Down vs. Maintain			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				Fuel Costs, Competition, and Demand				
Cost Savings	0.058***	0.050***	0.040***	0.032***				
	(0.007)	(0.006)	(0.007)	(0.006)				
Fuel Cost					0.010**	0.010*	0.005	-0.002
					(0.005)	(0.005)	(0.007)	(0.005)
Competitors	-0.044***	-0.054***	-0.038***	-0.026***	0.019**	0.018**	0.021***	0.012**
	(0.009)	(0.011)	(0.010)	(0.008)	(0.007)	(0.008)	(0.008)	(0.006)
Construction	0.050***	0.060***	0.050***	0.037***	-0.033***	-0.031***	-0.030***	-0.015*
	(0.010)	(0.012)	(0.011)	(0.010)	(0.011)	(0.011)	(0.011)	(0.009)
				Control Variables				
Kiln Age	0.016***	0.022***	0.019***	0.014***	0.016***	0.016***	0.016***	0.012***
	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Kiln Capacity	-0.997**	-0.601	-0.716	-0.420	-1.923***	-1.782***	-1.918***	-1.533***
	(0.355)	(0.441)	(0.447)	(0.341)	(0.416)	(0.457)	(0.471)	(0.361)
First Stage Residual	0.069***	0.083***	0.061***	0.047***	-0.020*	-0.022*	-0.023*	-0.004
	(0.014)	(0.016)	(0.017)	(0.014)	(0.011)	(0.013)	(0.013)	(0.010)
				Derived Statistics: Mean Elasticities of Pr(Adopt/Shut Down)				
WRT Cost Savings	0.818	0.713	0.565	0.686				
WRT Fuel Costs					0.381	0.386	0.204	-0.127
WRT Competitors	-1.769	-2.158	-1.570	-1.454	0.758	0.736	0.867	0.613
WRT Construction	1.302	1.542	1.300	1.160	-0.834	-0.817	-0.794	-0.446
				Specification Details				
Time Polynomial	no	1st Order	5th Order	no	no	1st Order	5th Order	no
Year Fixed Effects	no	no	no	yes	no	no	no	yes

Notes: The table summarizes results obtained from multinomial probit regressions in columns (i)-(iii) and a binomial probit regressions in column (iv). The sample is comprised of 8,192 kiln-year observations over 1973-2013. The dependent variable in Panel A is an indicator that equals one if the kiln is replaced with precalciner technology. The dependent variable in Panel B is an indicator that equals one if the kiln is shut down without replacement. The excluded instrument in the first stage is a 20-year lag on the number of nearby competitors. The elasticities of the estimated adoption probability with respect to *Cost Savings*, *Fuel Costs*, *Competitors*, and *Construction* are calculated for each observation and summarized with the mean. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

The Panel A results suggest that precalciner adoption is more likely when the cost savings are large, the number of nearby competitors is small, and there is more nearby construction activity. These effects are statistically significant and robust across the four columns. To evaluate magnitudes, we calculate the elasticity of the adoption probability with respect to the variables of interest, separately for each kiln-year observation in the data. The mean elasticity with respect to cost savings ranges over from 0.57 to 0.82; the mean elasticity with respect to the number of nearby competitors ranges from −1.45 to −2.16, and the mean elasticity with respect to nearby construction activity ranges from 1.16 to 1.54. Older kilns and smaller kilns are more likely to be replaced with a precalciner kiln, although the latter effect is not statistically significant with more flexible time controls. Finally, the first stage residual that controls for unobserved (but favorable) demand and cost conditions is positive and statistically significant, which is consistent with expectations and suggestive that the IV strategy is important.

The Panel B results indicate that kiln shutdown (without replacement) is more likely when fuel costs are high, the number of nearby competitors is large, and there is less nearby construction activity. These effects are statistically significant, with the exception of fuel costs if flexible time controls are employed (i.e., columns (iii) and (iv)). This latter result points to a data limitation: there is too little cross-sectional variation to identify the role of fuel costs in driving shutdown decisions once time-series variation is removed. The mean elasticity of the kiln shutdown probability with respect to the number of nearby competitors ranges from 0.61 to 0.87, and the mean elasticity with respect to construction activity ranges from −0.45 to −0.83. Older kilns and smaller kilns are more likely to be shut down. The first stage residual is negative – again consistent with it capturing favorable conditions – though its statistical significance is somewhat weaker in comparison to the adoption equation.

Figure 5 illustrates how the probabilities of adoption (Panel A) and shutdown (Panel B) change from respective one standard deviation increases in *Cost Savings* and *Fuel Cost*. Both of these probabilities depend upon the level of competition, which we highlight by plotting the effects over the range of *Competitors*. Panel A shows that amplified cost savings (e.g., via cap-and-trade regulation) increases the probability of adoption but only if competition is not too great: A one standard deviation in *Cost Savings* increases the adoption probability of a monopolist by more than three percentage points – a relatively large effect given that the empirical probability of adoption is only 1.8% in a given year. By contrast, the probability of adoption by a firm facing 30 nearby competitors (i.e., the 90th percentile) is negligible. Panel B suggests that if fuel costs are higher (again, e.g., via cap-and-trade regulation) then firms with many competitors may elect to shut down their kilns rather than adopt new

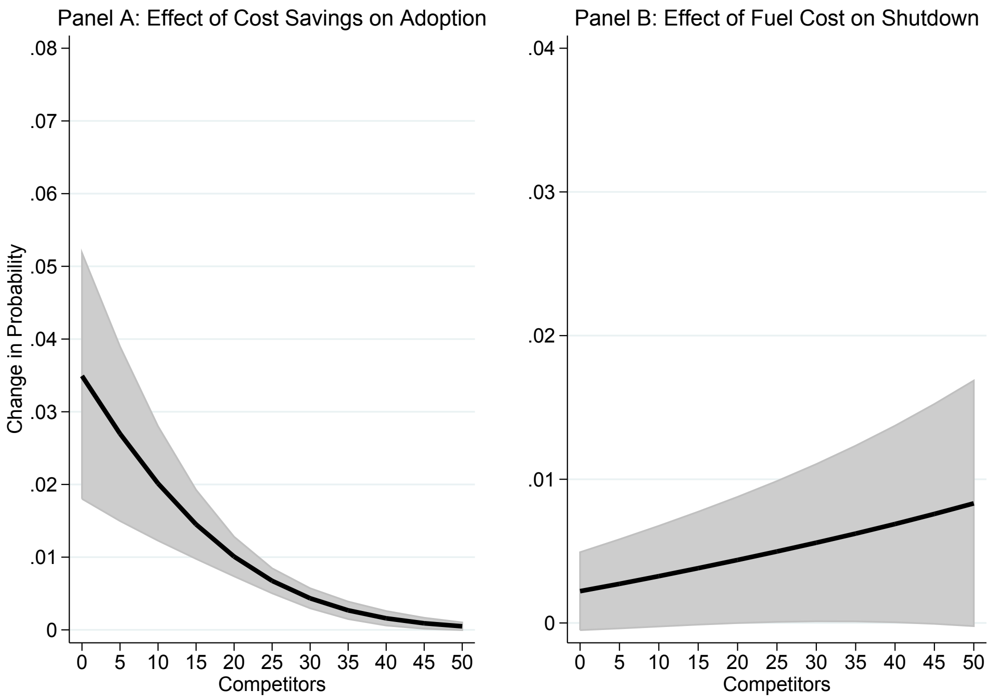

Figure 5: Changes in the Probabilities of Adoption and Shutdown

Notes: Panel A plots the change the adoption probability due to a one standard deviation increase in *Cost Savings* (which is 6.62). Panel B plots the change in the shutdown probability due to a one standard deviation increase in *Fuel Cost* (which is 9.63). Both effects are evaluated over the support of *Competitors*. Results are based on specification (ii) in Table 5. Other explanatory variables are held at their mean. The shaded regions provide 95% confidence intervals.

technology, however, this interactive effect does not appear to be statistically significant in the baseline specifications.

That the effect of *Cost Savings* depends on *Competitors* in the baseline specifications is due to the probit functional form. (Having many nearby competitors makes adoption sufficiently unlikely that there is little scope for other variables to matter.) The motivating theory goes somewhat further: competition should damp the benefit of technology adoption; not just the likelihood of adoption. To explore this possibility, we add an interaction of *Cost Savings* and *Competitors* in the adoption equation and an interaction of *Fuel Cost* and *Competitors* in the shutdown equation. We also interact the first stage residuals that control for unobserved demand and cost considerations.

Table 6 summarizes the results obtained with this augmented specification. The interaction of *Cost Savings* and *Competitors* is negative and statistically significant in columns (i)-(iii), providing additional support for the motivating theory. *Competitors* retains its level

Table 6: Probit Regression Results with Interaction

	Adopt Precalciner Technology vs. Maintain Old Kiln			
	(i)	(ii)	(iii)	(iv)
Fuel Cost Savings, Competition, and Demand				
Cost Savings	0.082***	0.079***	0.069***	0.045***
	(0.015)	(0.016)	(0.013)	(0.011)
Competitors	-0.027**	-0.034**	-0.013	-0.014
	(0.011)	(0.013)	(0.013)	(0.011)
Cost Savings	-0.0014**	-0.0015**	-0.0017***	-0.0008
×Competitors	(0.0006)	(0.0006)	(0.0006)	(0.0005)
Construction	0.045***	0.053***	0.038***	0.032***
	(0.010)	(0.012)	(0.012)	(0.010)
Derived Statistics: Mean Elasticities of Pr(Adoption)				
WRT Cost Savings	0.750	0.659	0.447	0.548
WRT Competitors	-1.548	-1.858	-1.119	-1.196
WRT Construction	1.169	1.356	1.012	1.020
Specification Details				
Control Variables	yes	yes	yes	yes
Time Polynomial	no	1st Order	5th Order	no
Year Fixed Effects	no	no	no	yes

Notes: The table summarizes results obtained from multinomial probit regressions in columns (i)-(iii) and a binomial probit regression in column (iv). The sample is composed of 8,192 kiln-year observations over 1973-2013. The dependent variable is an indicator that equals one if the kiln is replaced with precalciner technology. The excluded instrument in the first stage is a 20-year lag on the number of nearby competitors. The control variables include kiln age, kiln capacity, the first stage residual, and the first stage residual interacted with cost savings. The elasticities of the estimated adoption probability with respect to *Cost Savings*, *Competitors*, and *Construction* are calculated for each observation and summarized with the mean. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

effect in columns (i) and (ii), but this result dissipates with more flexible time controls and neither it nor the interaction achieve statistical significance with year fixed effects. The coefficients of other variables do not vary, however, and the mean elasticities are similar to those of the baseline specifications. The coefficients obtained for the shutdown equation (omitted for brevity) take the expected signs but are not statistically significant. As we develop below, however, some joint significance is obtained in columns (i) and (ii).

Figure 6 revisits the probabilities of adoption and shutdown with the augmented specification. Panel A again shows that if cost savings are amplified then adoption is more likely, provided that competition is not too great. The magnitude of the effects are somewhat larger than with the baseline specification, and the relationship is derived from the data rather than

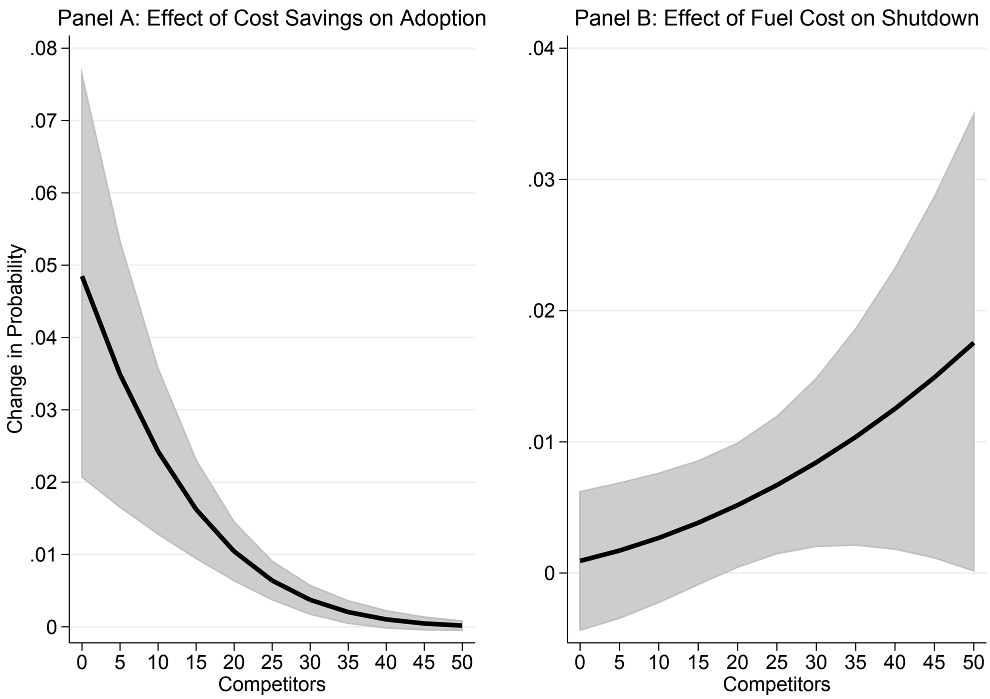

Figure 6: Changes in the Probabilities of Adoption and Shutdown with Interactive Effects
Notes: Panel A plots the change the adoption probability due to a one standard deviation increase in *Cost Savings* (which is 6.62). Panel B plots the change in the shutdown probability due to a one standard deviation increase in *Fuel Cost* (which is 9.63). Both effects are evaluated over the support of *Competitors*. Results are based on specification (ii) in Table 6. Other explanatory variables are held at their mean. The shaded regions provide 95% confidence intervals.

solely from the probit functional form. Panel B indicates firms that shut down their kilns in response to higher fuel costs tend to have more competitors, and this interactive effect is more apparent than in the baseline specifications. Further, the magnitudes are sizeable given that the empirical probability of shutdown is only 3.0% per year.

6.2 Robustness analysis

Table 7 evaluates robustness with respect to variable definitions and sample periods. We use multinomial probit regressions, but for brevity report results only for the adoption equation. A fifth order polynomial controls for time effects. Column (i) adds two alternative measures of the cost savings variable, based on fossil fuel prices five years ahead and behind the year of the observation. The alternative measures as shown do not correlate with adoption, but supports the timing assumptions that are implicit in the baseline variables – adoption does

Table 7: Probit Regression Results with Alternative Regressors

	Adopt Precalciner Technology vs. Maintain Old Kiln				
	(i)	(ii)	(iii)	(iv)	(v)
	Fuel Cost Savings, Competition, and Demand				
Cost Savings	0.032***	0.031***	0.031***	0.029***	0.039**
	(0.007)	(0.005)	(0.005)	(0.005)	(0.017)
Cost Savings ($t + 5$)	0.001				
	(0.005)				
Cost Savings ($t - 5$)	-0.007				
	(0.008)				
Competitors ($d < 400$)	-0.022***		-0.016**	-0.029**	-0.031**
	(0.006)		(0.006)	(0.012)	(0.015)
Competitors ($d < 200$)		-0.049***	-0.041***		
		(0.013)	(0.013)		
Construction	0.021***	0.030***	0.043***	0.042**	0.033**
	(0.007)	(0.007)	(0.013)	(0.018)	(0.013)
	Specification Details				
Control Variables	yes	yes	yes	yes	yes
Time Polynomial	5th Order	5th Order	5th Order	5th Order	5th Order
Sample Period	1973-2013	1973-2013	1973-2013	1973-1990	1991-2013

Notes: The table summarizes results obtained from multinomial probit regressions in columns (i)-(iii) and a binomial probit regression in column (iv). The sample is composed of 8,192 kiln-year observations over 1973-2013. The dependent variable is an indicator that equals one if the kiln is replaced with precalciner technology. The excluded instrument in the first stage is a 20-year lag on the number of nearby competitors. The control variables include kiln age, kiln capacity, and the first stage residual. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

not seem predetermined based on historical conditions nor is it driven in anticipation of future fossil fuel prices. Column (ii) modifies the number of competitors using a shorter distance radius of 200, but does not affect the results much. Column (iii) uses both the baseline radius (400) and the alternative radius (200), and both variables are found negative and statistically significant. The total effect of competition within a radius of 200 is $-0.016 - 0.041 = -0.057$, which suggests closer competitors matter more and is consistent with the role of transportation costs in creating localized market power. Columns (iv) and (v) use subsamples that respectively span 1973-1990 and 1991-2013. The results are strongly similar to those in the other columns, suggesting minimal differences in early versus late time periods of the regression sample.

Table 8 explores the performance of the IV strategy, again reporting results only for the adoption equation and using a fifth order polynomial control for time effects. Column (i)

Table 8: Probit Regression Results with Alternative IVs

	Adopt Precalciner Technology vs. Maintain Old Kiln			
	(i)	(ii)	(iii)	(iv)
	Fuel Cost Savings, Competition, and Demand			
Cost Savings	0.038***	0.038***	0.039***	0.039***
	(0.006)	(0.006)	(0.006)	(0.007)
Competitors	-0.018**	-0.022***	-0.028***	-0.034***
	(0.008)	(0.008)	(0.009)	(0.009)
Construction	0.033***	0.038***	0.042***	0.046***
	(0.010)	(0.011)	(0.011)	(0.011)
	Derived Statistics: Mean Elasticities of Pr(Adoption)			
WRT Cost Savings	0.532	0.539	0.552	0.560
WRT Competitors	-0.731	-0.935	-1.174	-1.420
WRT Construction	0.838	0.990	1.110	1.220
	Specification Details			
Time Polynomial	5th Order	5th Order	5th Order	5th Order
IV Lag Structure	No IV	5 Year	10 Year	15 Year

Notes: The table summarizes results obtained from multinomial probit regressions in columns (i)-(iii) and a binomial probit regression in column (iv). The sample is composed of 8,192 kiln-year observations over 1973-2013. The dependent variable is an indicator that equals one if the kiln is replaced with precalciner technology. The control variables include kiln age, kiln capacity, and the first stage residual. The elasticities of the estimated adoption probability with respect to *Cost Savings*, *Competitors*, and *Construction* are calculated for each observation and summarized with the mean. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

excludes the first stage residual, so that the endogeneity of competition is not addressed. The number of nearby competitors is still negatively associated with adoption, but the magnitude of the effect is reduced (e.g., the relevant mean elasticity falls from -1.57 to -0.73). The direction of this change is consistent with expectations given the source of bias. Columns (ii)-(iv) respectively use 5-year, 10-year, and 15-year lags on competition as the excluded instrument, instead of the 20-year lags used in the baseline specifications. These shorter lags produce smaller corrections for endogeneity, which suggests that some autocorrelation is present in the structural error term. However, the 15-year lag produces results that are quite similar to the 20-year lag: the mean elasticities of adoption with respect to competition are -1.42 and -1.57, and the difference is not statistically significant. This result provides support for the 20-year lag as a valid instrument. We further note that the 20-year lag on competition has great power in the first stage regressions: the F-statistic in the baseline specifications ranges from 2,474 to 3,417.

Table 9: Results from Alternative Empirical Models

Adopt Precalciner Technology vs. Maintain Old Kiln	Binomial Probit	Linear Probability	Hazard Rate
Fuel Cost Savings, Competition, and Demand			
Cost Savings	0.028***	0.0020***	0.068***
	(0.005)	(0.0006)	(0.007)
Competitors	-0.028***	-0.0012***	-0.043**
	(0.007)	(0.0003)	(0.018)
Construction	0.037***	0.0017***	0.038***
	(0.008)	(0.0004)	(0.021)
Derived Statistics: Mean Elasticities of Pr(Adoption)			
WRT Cost Savings	0.58	0.89	N/A
WRT Competitors	-1.91	-1.36	N/A
WRT Construction	1.47	1.21	N/A

Notes: The table summarizes results obtained from three different empirical models. The binomial probit and hazard rate models are estimated with maximum likelihood. The linear probability model is estimated with OLS. In the hazard rate model, a kiln shutdown is treated as a "competing risk." The sample is composed of 8,192 kiln-year observations over 1973-2013. The dependent variable is an indicator that equals one if the kiln is replaced with precalciner technology. All regressions incorporate control variables and a fifth order polynomial in time. The control variables are kiln age, kiln capacity, and the first stage residual. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

Finally, we explore whether the main results are driven by the use of the multinomial probit model. In successive columns of Table 9, we estimate the probability of adoption using a binomial probit, the linear probability model, and the "competing risks" semiparametric hazard rate model of Fine and Gray (1999). In the last model, the possibility of a kiln shutdown is incorporated into the model, but as an exogenous event rather than as a decision driven by particular economic circumstances. Each specification includes control variables (not shown) and a fifth order polynomial time control. The coefficients have similar signs, magnitudes, and statistical significance as those of the baseline regressions. Further, these results are robust to alternative treatments of time controls, with the exception that the hazard rate model does not converge with year fixed effects. We conclude that the multinomial probit regressions are picking up underlying empirical variation in the data and the results are not driven by specific functional forms.

6.3 Preemption incentives and the inverted-U

By creating incentives for preemption, competition could speed adoption in the medium even as it limits adoption in the long run. We examine whether this possibility is relevant in the data, based on the logic that if such preemption incentives are important then the presence of nearby precalciners should discourage adoption and/or encourage shutdown. Accordingly, we add the number of precalciner competitors within a radius of 400 to the baseline specification. We use two first-stage regressions and include 20-year lags on the number of competitors and the number of precalciner competitors as excluded instruments (both of which have considerable power). Both first-stage residuals are used as controls in the second-stage regressions. Table 10 indicates that number of nearby competitors retains its negative effect on precalciner adoption and its positive effect on shutdown decisions. The coefficients on nearby precalciner competitors are smaller and imprecisely estimated. Their signs also depend on the time controls. Thus, the data do not provide robust support for the hypothesis that precalciner competitors have a special importance in deterring technology adoption or hastening kiln shut downs.[12]

The motivating theory does not predict an inverted-U relationship in which adoption incentives are maximized with intermediate levels of competition. The inverted-U arises in some other models (e.g., Aghion et al (2005)), however, and we test for the relationship by allowing the benefits of adoption to have a quadratic relationship with the competition measure. The results are provided in Table 11. The quadratic term is found positive and statistically significant, which indicates that no empirical support for the inverted-U is found in this context. The net effect of *Competition* and *Competition*^{2} is negative in the support of the data, so the results should not be misinterpreted as implying that a sufficiently high degree of competition increases the benefits of competition. The analysis again provides corroborating support for the motivating theory and the underlying mechanism.

7 Conclusion

The research discussed herein explores the determinants of technology adoption in the portland cement industry. The technology in question, the precalciner kiln, reduces the marginal

[12] Another approach to testing for preemption is to see whether adoption is most likely for moderate levels of demand (e.g., Dafny (2005) Ellison and Ellison (2011)). The logic is that perhaps competitors definitely would not adopt with low enough demand and definitely would adopt with high enough demand, which isolates intermediate ranges of demand as possibly supporting preemption. We implement be adding a quadratic in *Construction*. The quadratic term has a t-statistic near zero in each of the baseline regressions, which again does not provide support for the importance of preemption in the data.

Table 10: Preemption Incentives

Panel A: Adopt Precalciner vs. Maintain Old Kiln				
	(i)	(ii)	(iii)	(iv)
Cost Savings	0.057***	0.051***	0.040***	0.031***
	(0.008)	(0.007)	(0.007)	(0.006)
Competitors	-0.045***	-0.055***	-0.040***	-0.029***
	(0.009)	(0.013)	(0.011)	(0.010)
Precalciner Competitors	-0.027	0.033	-0.010	0.016
	(0.039)	(0.063)	(0.060)	(0.057)
Construction	0.053***	0.060***	0.052***	0.041***
	(0.010)	(0.012)	(0.012)	(0.012)
Panel B: Shut Down vs. Maintain Old Kiln				
	(i)	(ii)	(iii)	(iv)
Fuel Cost	0.003	0.004	0.004	-0.003
	(0.007)	(0.007)	(0.007)	(0.006)
Competitors	0.022***	0.028**	0.022**	0.010
	(0.008)	(0.011)	(0.009)	(0.008)
Precalciner Competitors	-0.059	-0.090	0.008	0.043
	(0.045)	(0.068)	(0.062)	(0.058)
Construction	-0.036***	-0.040***	-0.031***	-0.012
	(0.011)	(0.013)	(0.012)	(0.011)
Specification Details (Both Panels)				
	(i)	(ii)	(iii)	(iv)
Control Variables	yes	yes	yes	yes
Time Polynomial	no	1st Order	5th Order	no
Year Fixed Effects	no	no	no	yes

Notes: The table summarizes results obtained from multinomial probit regressions in columns (i)-(iii) and a binomial probit regressions in column (iv). The sample is composed of 8,192 kiln-year observations over 1973-2013. The dependent variable in Panel A is an indicator that equals one if the kiln is replaced with precalciner technology. The dependent variable in Panel B is an indicator that equals one if the kiln is shut down without being replaced. There are two first stage regressions, for the number of competitors and the number of precalciner competitors, respectively. The excluded instruments are 20-year lags on the competition variables. The control variables include kiln age, kiln capacity, and both first stage residuals. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

Table 11: Tests for the Inverted-U

	Adopt Precalciner vs. Maintain Old Kiln			
	(i)	(ii)	(iii)	(iv)
	Fuel Cost Savings, Competition, and Demand			
Cost Savings	0.058***	0.049***	0.040***	0.031***
	(0.007)	(0.006)	(0.007)	(0.006)
Competitors	-0.065***	-0.067***	-0.074***	-0.049***
	(0.013)	(0.014)	(0.014)	(0.013)
Competitors^2	0.0004**	0.0003**	0.0007***	0.0004**
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Construction	0.053***	0.054***	0.049***	0.040***
	(0.010)	(0.012)	(0.012)	(0.010)
	Specification Details			
Control Variables	yes	yes	yes	yes
Time Polynomial	no	1st Order	5th Order	no
Year Fixed Effects	no	no	no	yes

Notes: The table summarizes results obtained from multinomial probit regressions in columns (i)-(iii) and a binomial probit regression in column (iv). The sample is composed of 8,192 kiln-year observations over 1973-2013. The dependent variable is an indicator that equals one if the kiln is replaced with precalciner technology. The excluded instrument is a 20-year lags on the competition variable. The control variables include kiln age, kiln capacity, and the first stage residual. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

costs of production. The empirical results suggest that adoption is more likely if the cost savings are large, and less likely if there are many nearby competitors. The results also suggest that competition damps the positive effect of cost savings on technology adoption. These results are consistent with the predictions of a two-stage theoretical model in which firms consider adoption, and then compete in accordance with Nash-Cournot competition. The mechanism is simple: competition denies firms the scale required to recoup the investment costs of technology adoption. The results thus support the relevance of the Schumpeterian hypotheses regarding firm size and innovation, within a specific market environment.

We conclude with a brief discussion of two possible research extensions. First, our empirical results suggest that the dynamic adjustment path of the cement industry in response to market-based regulation of CO_2 likely would involve some combination of investment and exit. Modeling this adjustment would be an interesting exercise. While some aspects of this thought experiment have been studied (e.g., Fowlie, Reguant and Ryan (2016)) additional progress could be made. New results could have bearing on the welfare consequences of regulation, and how market participants could most appropriately be compensated for losses. Second, when evaluated together with other recent research, our results are suggestive that preemption incentives are weaker in the presence of many competitors. However, the conditions under which preemption speeds technology adoption remain under-explored in the empirical literature. Further research could clarify the medium-run and long-run relationships between competition and technology adoption.

References

Aghion, Philippe and Jean Tirole, "The Management of Innovation," *Quarterly Journal of Economics*, 1994, *109*, 1185–1209.

____ **and Rachel Griffith**, *Competition and Growth: Reconciling Theory and Evidence*, The MIT Press, 2005.

____, **Antoine Dechezlepretre, David Hemous, Ralf Martin, and John Van Reenen**, "Carbon Taxes, Path Dependency and Directed Technical Change: Evidence from the Auto Industry," 2012. NBER Working Paper 18596.

____, **Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt**, "Competition and Innovation: An Inverted U-Relationship," *Quarterly Journal of Economics*, 2005, *120*, 701728.

Akhavein, Jalal, W. Scott Frame, and Lawrence J. White, "The Diffusion of Financial Innovations: An Examination of the Adoption of Small Business Credit Scoring in Large Banking Organizations," *Journal of Business*, 2005, *78* (2), 577–596.

Amemiya, Takesh, "The Estimation of a Simultaneous Equation Generalized Probit Model," *Econometrica*, 1978, *46*, 1193–1205.

Arrow, Kenneth J., "Economic Welfare and the Allocation of Resouces for Invention," 1962. In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, ed. J. Neyman, 507-532. Berkeley: University of California Press.

Bajari, Patrick, Han Hong, and Stephen Ryan, "Identification and Estimation of Discrete Games of Complete Information," *Econometrica*, 2010, *78* (5), 1529–1568. Working paper, Duke University.

Baldwin, William L. and John T. Scott, "Market Structure and Technological Change," 1987. In Lesourne and Sonnenschein (Eds.), *Fundamentals of Pure and Applied Economics*. Chur, Switzerland, and London: Harwook Academic Publishers.

Berry, Steven, "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 1992, *60*, 889–917.

Bresnahan, Timothy F. and Peter C. Reiss, "Entry and Competition in Concentrated Industries," *Journal of Political Economy*, 1991, *99* (5), 977–1009.

Busse, Meghan R., Christopher R. Knittel, and Florian Zettelmeyer, "Are Consumers Myopic? Evidence from New and Used Car Purchases," *American Economic Review*, 2013, *103* (1), 220–256.

Chicu, Mark, "Dynamic Investment and Deterrence in the U.S. Cement Industry," 2012. Job Market Paper, Northwestern University.

Ciliberto, Federico and Elie Tamer, "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 2009, *77*, 1791–1828.

Cohen, Wesley M., "Empirical Studies of Innovative Activity," 1995. In Stoneman, P. (Ed.), *Handbook of the Economics of Innovation and Technical Change*. Oxford: Basil Blackwell.

——, "Fifty Years of Empirical Studies of Innovative Activity and Performance," 2010. In Arrow, K.J. and Intriligator, M.D. (eds), *Handbook of the Economics of Innovation*, Volume 01. Elsevier.

—— **and Richard C. Levin**, "Empirical Studies of Innovation and Market Structure," 1989. In Schmalensee, R. and Willig, R. (Eds.), *Handbook of Industrial Organization*. Amsterdam: North-Holland.

Dafny, Leemore, "Games that Hospitals Play: Entry Deterrence in Hospital Procedure Markets," *Journal of Economics & Management Strategy*, 2005, *14* (3), 513–542.

Dasgupta, Partha and Joseph E. Stiglitz, "Industrial Structure and the Nature of Innovative Activity," *The Economic Journal*, 1980, *90* (358), 266–293.

Ellison, Glenn and Sara Fisher Ellison, "Strategic Entry Deterrence and the Behavior of Pharmaceutical IIncumbent Prior to Patent Expiration," *American Economic Journal: Microeconomics*, 2011, *3* (1), 1–36.

Fine, J. P. and R. J. Gray, "A Proportional Hazards Model for the Subdistribution of a Competing Risk," *Journal of the American Statistical Association*, 1999, *94*, 496–509.

Fowlie, Meredith, Mar Reguant, and Stephen P. Ryan, "Market-Based Emissions Regulation and Industry Dynamics," *Journal of Political Economy*, 2016, *124* (1), 249–302.

Ganapati, Sharat, Joseph S. Shapiro, and Reed Walker, "Energy Prices, Pass-Through, and Incidence in U.S. Manufacturing," 2016.

Genesove, David, "The Adoption of Offset Presses in the Daily Newspaper Industry in the United States," 1999. NBER Working Paper No. 7076.

Gil, Ricard, Jean-François. Houde, and Yuya Takahashi, "Preemptive Entry and Technology Diffusion in the Market for Drive-In Theaters," 2015. Working Paper.

Gilbert, Richard, "Looking for Mr. Schumpeter: Where Are We in the Competition-Innovation Debate?," 2006. In Jaffe, A.B., Lerner, J. and Stern, S. (Eds.), *Innovation Policy and the Economy*. Boston: The MIT Press, 159-215.

—— **and David Newbery**, "Preemptive Patenting and the Persistence of Monopoly," *American Economic Review*, 1982, *72*, 514–526.

Goettler, Ronald and Brett Gordon, "Does AMD Spur Intel to Innovate More?," *Journal of Political Economy*, 2011, *119* (6), 1141–1200. Mimeo.

Gowrisankaran, Gautum and Joanna Stavins, "Network Externalities and Technology Adoption: Lessons from Electronic Payments," *RAND Journal of Economics*, 2004, *35* (2), 260–276.

Griliches, Zvi, "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 1957, *25* (4), 501–522.

Hanlon, W. Walker, "Necessity is the Mother of Invention: Input Supplies and Directed Technical Change," 2014.

Hannan, Timothy H. and John M. McDowell, "The Determinants of Technology Adoption: The Case of the Banking Firm," *RAND Journal of Economics*, 1984, *15* (3), 328–335.

Igami, Mitsuru, "Estimating the Innovator's Dilemma: Structural Analysis of Creative Destruction in the Hard Disk Drive Industry, 1981-1998," *Journal of Political Economy*, 2015. Forthcoming.

—— **and Kosuke Uetake**, "Mergers, Innovation, and Entry-Exit Dynamics: The Consolidation of the Hard Disk Drive Industry, 1996-2015," 2016.

Kamien, Morton I. and Nancy L. Schwartz, *Market Structure and Innovation*, Cambridge: Cambridge University Press, 1982.

Karshenas, Massoud and Paul L. Stoneman, "Rank, Stock, Order, and Epidemic Effects in the Diffusion of New Process Technologies: An Empirical Model," *RAND Journal of Economics*, 1993, *24* (4), 203–528.

Langer, Ashley and Nathan H. Miller, "Automakers' Short-Run Responses to Changing Gasoline Prices," *Review of Economics and Statistics*, 2013, *95* (4), 1198–1211.

Linn, Joshua, "Energy Prices and the Adoption of Energy-Saving Technology," *The Economic Journal*, 2008, *118*, 1986–2012.

Mela, Carl F. and Praveen K. Kopalle, "The Impact of Collinearity on Regression Analysis: The Asymmetric Effect of Negative and Positive Correlations," *Applied Economics*, 2002, *34*, 667–677.

Miller, Nathan H. and Matthew Osborne, "Spatial Differentiation and Price Discrimination in the Cement Industry: Evidence from a Structural Model," *RAND Journal of Economics*, 2014, *45* (2), 221–247.

——, ——, and Gloria Sheu, "Pass-Through in a Concentrated Industry: Empirical Evidence and Regulatory Implications," 2015.

Newell, Richard G., Adam B. Jaffe, and Robert N. Stavins, "The Induced Innovation Hypothesis and Energy-Saving Technological Change," *Quarterly Journal of Economics*, 1999, *114* (3), 941–975.

Newey, Whitney, "Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables," *Journal of Econometrics*, 1987, *36*, 231–250.

Oster, Sharon, "The Diffusion of Innovation among Steel Firms: The Basic Oxygen Furnace," *The Bell Journal of Economics*, 1982, *13* (1), 45–56.

Perez-Saiz, Hector, "Building New Plant or Entering by Acquisition? Firm Heterogeneity and Entry Barriers in the U.S. Cement Industry," *RAND Journal of Economics*, 2015, *46* (3), 625–649.

Popp, David, "Induced Innovation and Energy Prices," *American Economic Review*, 2002, *92* (1), 160–181.

Rivers, Douglas and Quang H. Vuong, "Limited Information Estimators and Exogeneity TTest for Simultaneous Probit Models," *Journal of Econometrics*, 1988, *39*, 347–366.

Rose, Nancy and Paul L. Joskow, "The Diffusion of New Technologies: Evidence from the Electric Utility Industry," *RAND Journal of Economics*, 1990, *21* (3), 354–373.

Ryan, Stephen, "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, 2012, *80* (3), 1019–1062.

Schmidt-Dengler, Philip, "The Timing of New Technology Adoption: The Case of MRI," 2006. Working Paper.

Shapiro, Carl, "Theories of Oligopoly Behavior," 1989. In *Handbook of Industrial Organization, volume 1*, Richard Schmalensee and Robert Willing (eds). Elsevier.

Toivanen, Otto and Michael Waterson, "Market Structure and Entry: Where's the Beef?," *RAND Journal of Economics*, 2005, *36* (3), 680–699.

Tombak, Mihkel M., "Mergers to Monopoly," *Journal of Economics & Management Strategy*, 2002, *11* (3), 513–546.

Van Oss, Hendrik G., "Background Facts and Issues Concerning Cement and Cement Data," 2005. Open-File Report 2005-1152, U.S. Department of the Interior, U.S. Geological Survey.

Vogt, William B., "Detecting Strategic Behavior in Technology Adoption: The Example of Magnetic Resonance Imaging," 2000. Unpublished working paper.

Appendices

A Measuring Fuel Costs

We calculate the energy requirements of production based on the labor-energy input surveys of the PCA. There is no discernible change in the requirements over 1990-2010, conditional on the kiln type. We calculate the average mBtu per metric tonne of clinker required in 1990, 2000, and 2010, and apply these averages over 1990-2013. These are 3.94, 4.11, 5.28, and 6.07 mBtu per metric tonne of clinker for precalciner kilns, preheater kilns, long dry kilns, and wet kilns, respectively. A recent USGS survey accords with our calculations (Van Oss (2005)). Technological improvements are evident over 1973-1990 within kiln type: in 1974, the energy requirements were 6.50 mBtu per metric tonne of clinker at dry kilns (a blended average across dry kiln types) and 7.93 mBtu per metric tonne of clinker at wet kilns. We assume that improvements are realized linearly over 1973-1990. We scale down by our calculated energy requirements by five percent to reflect that a small amount of gypsum is ground together with the kiln output (i.e., clinker) to form cement.

Plants sometimes list multiple primary fuels in the PIS. In those instances, we calculate fuel costs with the coal price if coal is among the primary fuels; otherwise, we use natural gas prices if natural gas is among the multiple fuels. We use oil prices only if oil is the only fossil fuel listed. In the 1980s, petroleum coke supplements or replaces coal at many kilns. The price of coal and petroleum coke are highly correlated, and we simply use the coal price for those observations. Figure A.1 plots fossil fuel prices and usage over the sample period. In the mid-1970s, coal and natural gas wer the most popular fuel choices, while only a small subset of plants used oil. Coal quickly came to dominate the industry due to a change in relative prices, and fuel costs thereafter track the coal price.

Our methodology does not incorporate secondary fuels, the most popular of which are waste fuels such as solvents and used tires. The labor-energy input surveys of the PCA indicate that waste fuels account for around 25% of the energy used in wet kilns and 5% of the energy used in dry kilns. We do not have data on the prices of waste fuels but understand them to be lower on a per-mBtu basis than those of fossil fuels. Accordingly, we construct an alternative fuel cost measure in which we scale down the fossil fuel requirements of wet and dry kilns in accordance with the survey data. Whether this adjustment better reflects the fuel costs of marginal output depends in part on (i) the relative prices of waste and fossil fuels and (ii) whether the average fuel mix reported in the survey data reflect the marginal fuel mix. On the latter point, if marginal clinker output is fired with fossil fuels then our

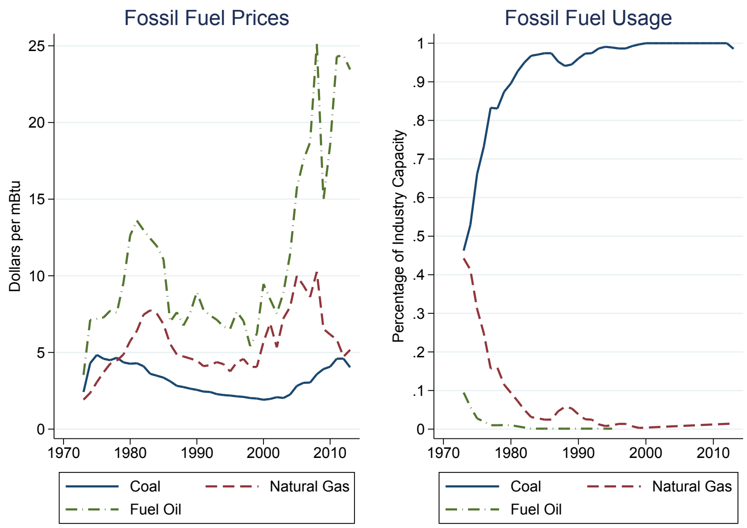

Figure A.1: Fossil Fuel Prices and Usage 1973-2013

baseline measurement should reflect marginal fuel costs more closely than the alternative measurement. Regardless, our regression results are not very sensitive to the adjustment.

B Theory

B.1 Focal firm model

What must be shown is that $\frac{\Delta b_i(\Delta c, c_i^0, \bar{\bar{c}}, a, N)}{\Delta N} < 0$ under condition (5), because it is evident from equation (4) that increasing N reduces the (positive) derivative of benefits with respect to Δc under condition (5). Benefits decrease in N if

$$b_i(\Delta c, c_i^0, \bar{\bar{c}}, a, N+1) < b_i(\Delta c, c_i^0, \bar{\bar{c}}, a, N)$$

Plugging in equations (3) and (4) yields

$$\frac{N+1}{N+2}\frac{a - \widehat{c}_i + (N+1)(\bar{\bar{c}} - \widehat{c}_i)}{N+2} < \frac{N}{N+1}\frac{a - \widehat{c}_i + N(\bar{\bar{c}} - \widehat{c}_i)}{N+1} \tag{B.1}$$

and with some effort this can be manipulated to obtain condition (5).

B.2 Long run equilibrium

Derivation of equation (9)**.** We solve for the SPE number of adopters, L^*, which is characterized by the inequalities $k \leq q_1(L = L^*)^2 - q_0(L^* - 1)^2$ and $k > q_1(L^* + 1)^2 - q_0(L^*)^2$. (We have condensed the notation for brevity.) We start with the first inequality. It can be reexpressed:

$$k \leq \big(q_1(L^*) + q_0(L^* - 1)\big) \times \big(q_1(L^*) - q_0(L^* - 1)\big) \tag{B.2}$$

because $x^2 - y^2 = (x + y)(x - y)$. It can be shown than $q_1(|L) = q_0(L) + \Delta c$ for any L. Thus, plugging in for q_1 yields

$$k \leq \big(q_0(L^*) + \Delta c + q_0(L^* - 1)\big) \times \big(q_0(L^*) + \Delta c - q_0(L^* - 1)\big)$$

If there are L firms with costs of c^1 and $N - L$ firms with costs of c^0, then equation (3) can be used to obtain the equilibrium quantity of a non-adopter:

$$q_0(L) = \frac{a - Nc^0 + Lc^1 + (N - L - 1)c^0}{N + 1} = \frac{a - c^0 - L\Delta c}{N + 1} \tag{B.3}$$

Substituting into the inequality yields

$$k \leq \left(2\frac{a - c_0}{N + 1} - \left(\frac{L}{N + 1} + \frac{L - 1}{N + 1}\right)\Delta c + \Delta c\right)\left(-\frac{L}{N + 1}\Delta c + \Delta c + \frac{L - 1}{N + 1}\Delta c\right)$$

Collecting terms,

$$k \leq \left(2\frac{a - c_0}{N + 1} + \frac{N - 2L}{N + 1}\Delta c\right)\left(\frac{N}{N + 1}\Delta c\right)$$

The second inequality that characterizes L^* is analogous, so we also have

$$k > \left(2\frac{a - c_0}{N + 1} + \frac{N - 2(L + 1)}{N + 1}\Delta c\right)\left(\frac{N}{N + 1}\Delta c\right) \tag{B.4}$$

Combining the inequalities yields

$$\frac{N - 2(L + 1)}{N + 1} < \frac{k}{(\Delta c)^2}\frac{N + 1}{N} - \left(2\frac{a - c_0}{\Delta c}\right)\left(\frac{1}{N + 1}\right) \leq \frac{N - 2L}{N + 1}$$

Multiplying by $N + 1$ and subtracting N yields

$$-2(L + 1) < \frac{k}{(\Delta c)^2}\frac{(N + 1)^2}{N} - 2\frac{a - c_0}{\Delta c} - N \leq -2L$$

Finally, dividing by negative two flips the direction of the inequalities

$$L \leq N + \frac{a - c_0}{\Delta c} - \frac{1}{2}\frac{k}{(\Delta c)^2}\frac{(N+1)^2}{N} < L + 1 \tag{B.5}$$

This is equation (9) in the text.

Proof of Result 3. If $k > 2(\Delta c)^2$, it can be shown that as long as not everyone adopts, then the number of adopters decreases with competition in some interval. To see this, we work with middle part of the inequality in equation (9) and take the derivative of the bound

$$N + \frac{a - c^0}{\Delta c} - \frac{1}{2}\frac{k}{(\Delta c)^2}\frac{(N+1)^2}{N},$$

which is

$$1 - \frac{1}{2}\frac{k}{(\Delta c)^2}\frac{N^2 - 1}{N^2}.$$

We want to know when this derivative is negative, equivalently, finding the zeros of this equation. Note that the zeros of this equation solve

$$N^2\left(1 - \frac{1}{2}\frac{k}{(\Delta c)^2}\right) + \frac{1}{2}\frac{k}{(\Delta c)^2} = 0.$$

The above equation defines a parabola in N that will slope upwards if $k < 2(\Delta c)^2$ and downwards if $k > 2(\Delta c)^2$. The parabola is maximized or minimized at $N = 0$. The roots of the equation are

$$\pm\sqrt{\frac{-\frac{1}{2}\frac{k}{(\Delta c)^2}}{1 - \frac{1}{2}\frac{k}{(\Delta c)^2}}}$$

If $k < 2(\Delta c)^2$ then this equation has no real roots. Thus the parabola slopes upwards, is always positive, and more competition always increases the incentive to adopt. If instead it slopes downwards, there is one real positive root which I'll call N_1. If $N > N_1$ and $L^* + 1 \leq N$, then it must be that increasing N weakly decreases the number of firms in the market (the inequality can't be strong due to the integer nature of the number of firms). Note that there will be some point at which the upper bound on L^* becomes negative, since the upper bound is decreasing in N, and the upper bound on L^*/N converges to a negative number as N goes to infinity. This means there is some number N_2 such that for $N > N_2$ nobody adopts. Within the interval N_1 and N_2, since L decreases as N rises, the fraction of adopters must also decrease. The bound N_1 is unnecessarily tight, as the fraction of firms that adopt could decrease in N even if L is increasing in N – it is sufficient that L increases in N at a slower rate than N.

C Preemption

Start with *ex ante* symmetric firms as in the two-stage model. Timing of the game:

1. Leader can adopt or not, for investment cost k_1.
2. Payoffs given by Cournot.
3. All firms can adopt or not, for capital cost $k_2 < k_1$. Decisions are sequential with leader moving last.
4. Payoff given by Cournot, or possibly a stream of payoffs.

Define some combinations of parameters for which $L^* = 2$ if $N = 2$ and $L^* = 1$ if $N = 3$.

Result 1: If $N = 2$, there exists a discount rate such that the leader foregos early adoption in order to take advantage of lower investment costs. Both firms adopt in the later stage.

Result 2: If $N = 3$, then there exists a discount rate for which the leader adopts in the early stage and no firms adopt after.

Proofs coming soon.

D Standard Error Adjustment

The two stage conditional maximum likelihood estimator developed by Rivers and Vuong (1988) and Newey (1987) was done in the context of the binomial, rather than the multinomial probit. However it is not difficult to extend the procedure proposed by Newey (1987) to the multinomial probit. We will demonstrate how to do this in the context of our model. In this appendix, we follow the notation used in Newey (1987) rather than what appears in the body of this paper. Denote the endogenous regressor as $y_{2,it}$, the left hand side variables as $y^A_{1,it}$ and $y^S_{1,it}$, the right hand side variables that are common to the first and both second step equations are $\boldsymbol{x}_{11,it}$, variables that are in the second step equations but not the first we denote at $\boldsymbol{x}^A_{12,it}$, $\boldsymbol{x}^S_{12,it}$, variables in the first step but not the second we denote as $\boldsymbol{x}_{2,it}$. We can write down the equations that define the two step estimator as

$$
\begin{aligned}
y_{1,it}^{A} &= y_{2,it}\beta^{A} + \boldsymbol{x}_{11,it}\boldsymbol{\gamma}_{1}^{A} + \boldsymbol{x}_{12,it}^{A}\boldsymbol{\gamma}_{2}^{A} + u_{it}^{A} \\
y_{1,it}^{S} &= y_{2,it}\beta^{S} + \boldsymbol{x}_{11,it}\boldsymbol{\gamma}_{1}^{S} + \boldsymbol{x}_{12,it}^{S}\boldsymbol{\gamma}_{2}^{S} + u_{it}^{S} \\
y_{2,it} &= \boldsymbol{x}_{11,it}\boldsymbol{\Pi}_{1} + \boldsymbol{x}_{2,it}\boldsymbol{\Pi}_{2} + v_{it}
\end{aligned}
$$

The procedure follows four steps:

1. In the first step, we estimate the coefficients $\hat{\boldsymbol{\Pi}}_1$ and $\hat{\boldsymbol{\Pi}}_2$ using regression, and recover the residuals $\hat{v}_{it}$.

2. In the second step, we estimate a reduced-form probit of the adopt/shutdown decisions on all the exogenous variables, and the fitted residuals from step 1. The estimation equations can be written as:

$$
\begin{aligned}
y_{1,it}^{A} &= \boldsymbol{w}_{it}^{A}\boldsymbol{\alpha}^{A} + \lambda^{A}\hat{v}_{it} + u_{it}^{A} \\
y_{1,it}^{S} &= \boldsymbol{w}_{it}^{S}\boldsymbol{\alpha}^{S} + \lambda^{S}\hat{v}_{it} + u_{it}^{S},
\end{aligned}
$$

   where we define $\boldsymbol{w}_{it}^{k} = [\boldsymbol{x}_{2,it} \quad \boldsymbol{x}_{11,it} \quad \boldsymbol{x}_{12,it}^{k}], \quad k \in \{A, S\}$. An important point here is that the reduced-form coefficients $\boldsymbol{\alpha} = [\boldsymbol{\alpha}^{A\prime} \quad \boldsymbol{\alpha}^{S\prime}]'$ can be written in terms of the structural coefficients as

$$
\boldsymbol{\alpha} = D(\boldsymbol{\Pi}_1, \boldsymbol{\Pi}_2)\boldsymbol{\delta},
$$

   where $\boldsymbol{\delta} = [\beta^{A\prime} \quad \gamma_1^{A\prime} \quad \gamma_2^{A\prime} \quad \beta^{S\prime} \quad \gamma_1^{S\prime} \quad \gamma_2^{S\prime}]'$, and the matrix $D(\boldsymbol{\Pi}_1, \boldsymbol{\Pi}_2)$ is defined as follows:

$$
D(\boldsymbol{\Pi}_1, \boldsymbol{\Pi}_2) = \begin{bmatrix}
\boldsymbol{\Pi}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\boldsymbol{\Pi}_1 & \boldsymbol{I}_1^A & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \boldsymbol{I}_2^A & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Pi}_2 & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Pi}_1 & \boldsymbol{I}_1^s & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I}_2^S
\end{bmatrix},
$$

   where the identity matrix $\boldsymbol{I}_j^k$ corresponds to the size of the $\boldsymbol{\gamma}_j^k$ parameter vector, for $j \in \{1, 2\}$ and $k \in \{A, S\}$. Amemiya (1978) shows that a consistent and asympotically efficient estimator for the structural parameters $\boldsymbol{\delta}$ can be derived from the minimization problem

$$\min_{\boldsymbol{\delta}} (\tilde{\boldsymbol{\alpha}} - \hat{D}\boldsymbol{\delta})' \hat{\boldsymbol{\Omega}}^{-1} (\tilde{\boldsymbol{\alpha}} - \hat{D}\boldsymbol{\delta}),$$

where $\tilde{\boldsymbol{\alpha}}$ and $\hat{D}$ are consistent estimates of $\boldsymbol{\alpha}$ and $D(\boldsymbol{\Pi}_1, \boldsymbol{\Pi}_2)$, respectively, and $\hat{\boldsymbol{\Omega}}$ is a consistent estimator of the variance of $\sqrt{N}(\tilde{\boldsymbol{\alpha}} - \hat{D}\boldsymbol{\delta})$. Under these conditions the distribution of the structural estimates $\hat{\boldsymbol{\delta}}$ are asymptotically normal with variance $(\hat{D}'\hat{\boldsymbol{\Omega}}^{-1}\hat{D})^{-1}$. We can construct $\hat{D}$ from our first stage estimates in step 1. Newey (1987) shows that the variance $\boldsymbol{\Omega}$ can be expressed in the following form:

$$\boldsymbol{J}_{\alpha\alpha}^{-1} + (\boldsymbol{\lambda} - \boldsymbol{\beta})' \boldsymbol{\Sigma}_{vv} (\boldsymbol{\lambda} - \boldsymbol{\beta}) \boldsymbol{Q}^{-1}. \tag{D.1}$$

The first term, $\boldsymbol{J}_{\alpha\alpha}^{-1}$, is the covariance matrix of the reduced-form probit parameters $\boldsymbol{\alpha}$. We describe how to construct an estimator of the second term in the fourth step.

3. In the third step, we estimate $\hat{\beta}^A$, $\hat{\beta}^S$, $\hat{\boldsymbol{\gamma}}_1^A$, $\boldsymbol{\gamma}_2^A$ using a multinomial probit. We include the residuals from stage 1 as regressors.

4. In the final step, we regress $y_{2,it}(\boldsymbol{\lambda} - \boldsymbol{\beta})$ on the reduced form regressors for each equation, $\boldsymbol{w}_{it}^A$ and $\boldsymbol{w}_{it}^S$. We run this as a Seemingly Unrelated Regressions system in two equations, and recover the variance matrix of the parameters. From Newey (1987) this variance matrix is an estimator for the second term of equation (D.1). Together with the estimated variance matrix of the reduced-form probit parameters, and the estimated $\hat{D}$, we can form an estimate of the variance of the structural parameters $\boldsymbol{\delta}$.