Market concentration in homebuilding

Jacob Cosman and Luis Quintero

Carey Business School
Johns Hopkins University

April 12, 2018

We investigate the impact of increasing concentration in local residential construction markets on housing cycle dynamics. We show that the increase in concentration has led to greater unit price volatility, less production, and fewer vacant unsold units. Our results imply that the greater concentration has decreased the annual value of new housing production by $144 billion. Because housing is a determinant of the business cycle these findings provide further evidence that the secular decline in competitive intensity in the American economy is altering macroeconomic dynamics[1].

1 Introduction

The housing market cycle is an integral leading component of the overall business cycle. As of 2016, housing consumption accounts for 16% of total personal consumption expenditures and 11% of GDP (Bureau of Economic Analysis, 2017b). Primary residential mortgages account for two-thirds of all household debt (Board of Governors of the Federal Reserve System, 2017). Housing is central to households' consumption and investment decisions (Hurst and Stafford, 2004) and household consumption during recessions is heavily influenced by their decline in housing wealth (Mian et al., 2013; Mian and Sufi, 2016). The housing market cycle is an important component of overall macroeconomic dynamics and in particular the transmission of financial shocks to the real economy (Guerrieri and Uhlig, 2016). Leamer

---

[1] We are grateful to Judith Ricks and Andrea Lamorgese as well as seminar participants at the 2017 International Industrial Organization Conference, the 2017 American Real Estate and Urban Economics Association National Conference, the 2017 Richmond Federal Reserve Regional Economic Workshop, the 2017 Urban Economics Association Meeting, George Washington University, and the USC Dornsife-INET Conference on Housing, Urban Development, and the Macroeconomy for their insightful comments. Partial funding for this research by The Black & Decker Research Fund of the Johns Hopkins Carey Business School is gratefully acknowledged.

(2007) finds that housing market cycles are typically driven by fluctuations in the volume of production rather than fluctuations in prices[2]. Accordingly, understanding the production of new housing is essential to understanding macroeconomic dynamics.

In the Great Recession and its aftermath the market structure in the production of new housing shifted. As discussed in detail below, local market concentration in residential construction increased drastically as national firms increased production. Home builders are aware of the lower competition and consider it beneficial to their business; a May 2017 Wall Street Journal article featured the headline "Fewer Home Builders Means Happier Home Builders" (Lahart, 2017).

This shift in the competitive environment impacts the dynamics of housing markets. As Mueller (1995) notes, firms competing to build quickly to satisfy unmet demand and gain a first-mover advantage in a growing market are the drivers of real estate market cycles:

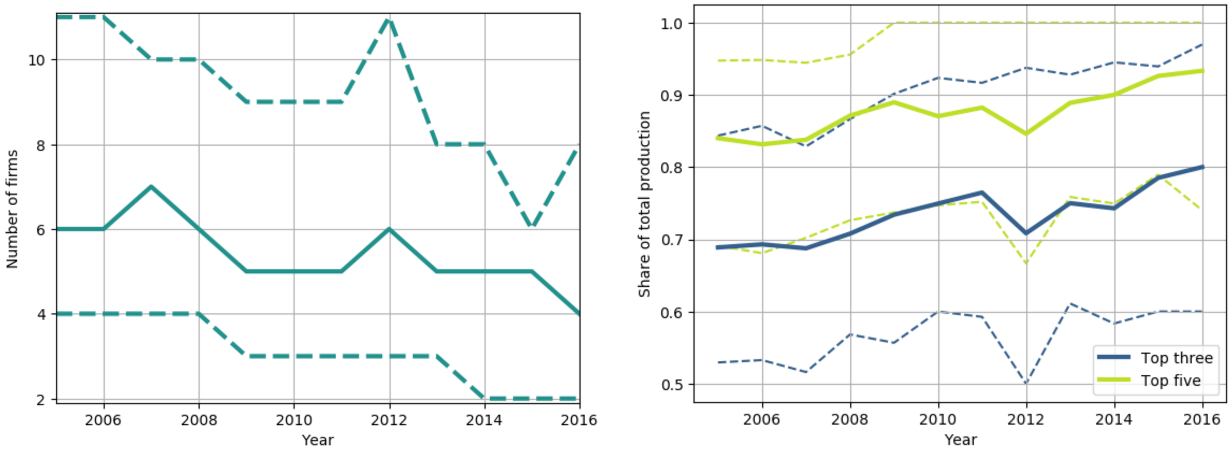
> In a competitive capitalistic market, developers must speculate and start the process of planning development or building new products earlier than the actual demand materializes to edge out other developers who also want a share of the market. In the absence of collusion, this speculative behavior, along with the lumpy nature of real estate product, makes it easy to overshoot actual needs.

The theoretical framework of Mueller (1995) broadly informs market participants' understanding of real estate cycle dynamics. Media reports similarly emphasize the role of competition in generating a rush to build in the commercial (Schnurman, 2010), retail (Sandler, 2000), and residential (O'Connell, 2011; Gopal, 2016) construction sectors. However, to date, the empirical literature has not investigated the role of competitive intensity in the production of build new housing.

In this study we investigate how competition between homebuilders drives the trajectory of housing cycles. We use a novel panel data set to demonstrate the effect of market concentration on housing market dynamics. Specifically, we show that the secular decline in competitive intensity over the last decade has led to fewer housing units built, a decline in the inventory of vacant unsold housing units, and greater price volatility. As reviewed in Piazzesi and Schneider (2016), studies in the macroeconomic literature often model variation in housing construction as attributable to highly volatile total factor productivity. Our micro-founded approach provides an alternative perspective on the dynamics of the industry.

---

[2] Leamer (2015) attributes the distinctive dynamics of the most recent housing market cycle to the specific confluence of monetary policy and mortgage securitization at the time.



(a) Number of firms accounting for 90% of housing construction. The solid line shows the median market and the dashed lines show the first and third quartiles.

(b) Share of production accounted for by largest three and largest five firms in each market. The solid line shows the median market and the dashed lines show the first and third quartiles.

Figure 1: Measures of high and increasing concentration in local housing markets. Section 3.2.1 describes the data set in detail.

1.1 Market concentration in housing construction

Housing production is highly concentrated in local markets. For example, in our sample[3], Baker Residential built 37% of all new housing units in Bayonne, NJ and Technical Olympic built 47% of all new housing units in Centreville, VA, between 2005 and 2016. Moreover, market concentration is rising over time. The Craftmark Group was responsible for 3% of new units in Annapolis, MD between 2005 and 2007 but 43% of new units between 2014 and 2016 and Baker Residential built no units in Middletown, NY between 2005 and 2007 but 37% of new units between 2014 and 2016.

Figure 1 shows the high concentration in the local housing markets in our sample. Over the sample period, the share of production by the largest firms in each market increased and the number of firms producing 90% of all new units decreased. By 2016, in the most concentrated quartile of all markets, two or fewer firms produced at least 90% of all new housing.

Figure 2 shows the distribution of Herfindahl indices across markets in the sample as of 2006 and in 2015. For regulatory purposes the United States Department of Justice and the Federal Trade Commission deem any market with a Herfindahl index between 1500 and

---

[3] As discussed in Section 3.2.1, our sample includes 132 housing markets in the northeastern United States.

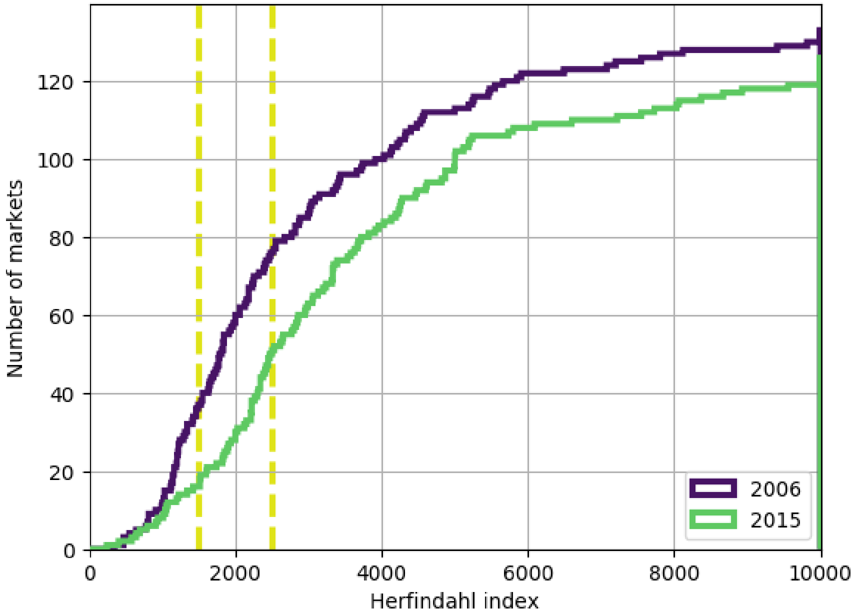


Figure 2: Cumulative distribution of Herfindahl indices for all markets in the sample in 2006 and in 2015. The dashed lines denotes the Federal Trade Commission standard for “moderately concentrated” and “highly concentrated” markets.

2500 to be “moderately concentrated” and a Herfindahl index in excess of 2500 to be “highly concentrated” (U.S. Department of Justice and Federal Trade Commission, 2010). As shown, the entire distribution of Herfindahl indices has shifted towards higher concentration during this period. By 2015 60% of markets surpassed the “highly concentrated” threshold. Martín and Whitlow (2012) note that this shift is a relatively new phenomenon from the 2000s onwards.

Three changes to the national environment have contributed to the increase in market concentration over this period:

1. Many homebuilding firms filed for bankruptcy in 2008 in the wake of the housing market downturn (Thompson, 2009). Highly active firms in our sample which filed for bankruptcy include Caruso Homes (Merle, 2008), Woodside Homes (Beebe, 2012), WCI Communities (Kessler, 2008), and Gemcraft (Mirabella, 2009). Bankruptcy limited these firms’ ability to construct new housing for several years.

2. A federal legislative stimulus measure late in 2009 increased the ability of homebuilders to include previous years’ losses in their taxes. Graham and Kim (2009) examine the impact measure on one large national firm; they find that “the liquidity effect of the

proposed increase in the carryback period appears substantial". This change in rules provided particular benefit to Lennar (Cook, 2010) as well as Toll Brothers, KB Homes, and other large national firms that had been active in markets most vulnerable to the housing market downturn (Barr, 2010). According to Corkery and Drucker (2009), the thirteen largest homebuilders expected \$2.4 billion of tax refunds in 2009 as a result of the rule change.

3. In recent years large national homebuilders in our data set have merged with other national homebuilders. In particular, Pulte Homes and Centex merged in 2009 to create (at the time) the largest homebuilding firm in the country (Clifford, 2009), Tri Point merged with Weyerhauser in 2013 (Sorkin, 2013), and Standard Pacific and Ryland merged to form CalAtlantic in 2015 (Hudson, 2015)[4].

These changes had economically meaningful impacts for the markets in our sample. For example, the Elliott Building Group was present in several markets in Pennsylvania and New Jersey but following its bankruptcy in 2007 (Crocker, 2007) it was no longer active in our sample. Conversely, the national-scale Pulte Homes accounted for a much larger share of housing construction in many markets after the introduction of federal stimulus. In Germantown, MD the Pulte Homes share of closings increased from 22% for the years from 2006 through 2008 to 48% for the years from 2010 through 2012 while in Odenton, MD its share increased from zero to 26% over the same period. Taken together, these changes to the competitive environment appear to have favoured the largest firms and led to much of the increase in concentration shown in Figure 1.

Financial markets also provide evidence for the success of large national firms. Figure 3 compares the performance of the largest exchange-traded fund based on homebuilder stocks (Google Finance API, 2018a) to the S&P 500 during the economic recovery (Google Finance API, 2018b). Large national firms in our sample comprise the majority of the fund's portfolio. Even during a historically long market expansion, large national homebuilding firms have outperformed the overall stock market.

Homebuilding firms and other industry participants recognize the advantages of being the dominant firm in a concentrated local market. Lahart (2017) reports that builders are more optimistic about their future success following a reduction in local competitive intensity.

---

[4] Consolidation has continued after the end of our sample. Specifically, in 2017 Lennar purchased WCI Communities (Lane, 2017b) and then merged with CalAtlantic (Bray and Goldstein, 2017) to form the largest homebuilding firm in the country (Builder Magazine M&A, 2017; Gara, 2017). In the same year DR Horton purchased Forestar Group at the culmination of a bidding war with Starwood Capital Group (Lane, 2017a)

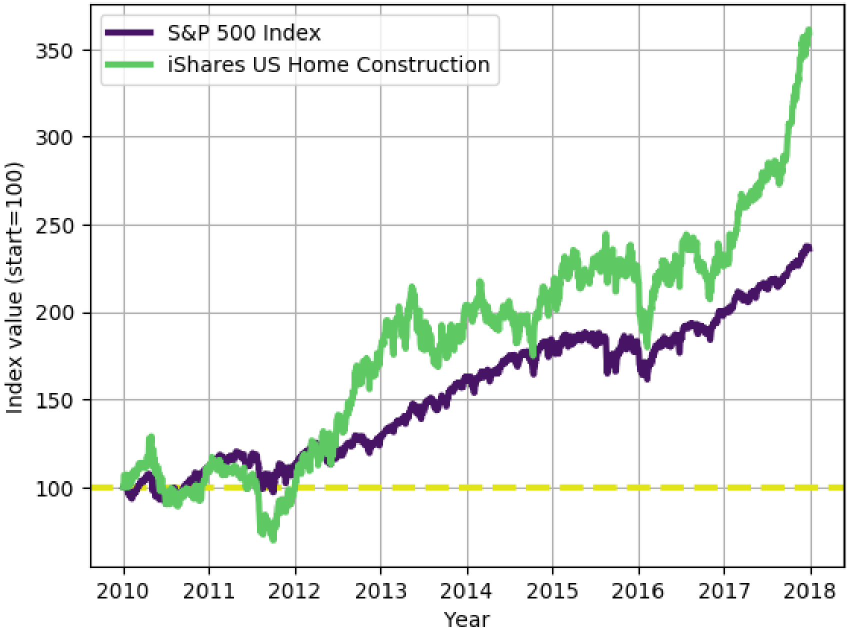


Figure 3: Comparison of the performance of S&P 500 index with iShares US Home Construction (the largest largest exchange-traded fund based on homebuilder stocks by market capitalization). Both series are normalized to 100 at the beginning of 2010.

Builder Magazine M&A (2017) cites an analyst asserting that the merger of Lennar and CalAtlantic to "dominate the housing market" in areas where both firms were active in explaining the benefits of the merger.

In addition to these strategic considerations larger homebuilding firms benefit from substantial production advantages relative to smaller firms. O'Hollaren (2017) enumerates several of these advantages including economies of scale, the ability to handle design and development in-house, the potential for joint ventures with government and industry, brand name recognition, and financing packages for consumers. Martín and Whitlow (2012) note that large firms also benefit from bulk purchases that lower the cost of materials, superior access to capital markets, and land inventories that allow for less costly production of new housing. Khouri (2015) reports that when Standard Pacific and Ryland Group merged to form CalAtlantic the combined firm was able to reduce its corporate staff by 10%. Lane (2017a) reports that the merger of Lennar and CalAtlantic would allow for greater efficiencies in purchasing materials and land as well as hiring labour. Porter (2003) suggests that larger firms' access to large volumes of patient capital through corporate bond markets and greater staff capabilities makes them better equipped to navigate local land use regulation.

According to Metcalf (2018), the complexity of local government land use approval processes constitutes a substantial fixed cost to homebuilders as well as a barrier to entry.

Given these strategic and cost advantages enjoyed by the largest homebuilding firms it seems likely that the current consolidation will persist and that many local markets will remain highly concentrated. McGraw Hill Construction (2006) predicts that “homebuilder profitability will favor large multi-regional players” while O’Hollaren (2017) notes that “revenue is increasingly concentrated among the largest businesses in the industry”. Accordingly, the role of high concentration in housing market cycle dynamics merits further investigation.

1.2 Literature review

In light of the integral role of housing in the macroeconomy, understanding firms’ strategic decisions in the production of new housing has implications for understanding not only urban growth but also business cycle dynamics. As discussed in Davila and Korinek (2016) the macroeconomic literature on these cycles has largely focused on financial frictions — particularly the roles of incomplete markets and collateral constraints. He and Kondor (2016) investigate how the impact of the financial frictions change over the business cycle to generate investment “waves”. In the financial literature, theoretical results including Grenadier (2002) indicate that industry competitiveness plays a role as firms in a more competitive industry will rush to take advantage of scarce investment opportunities. Aguerrevere (2003) provides theoretical evidence that in an oligopolistic industry where building new capacity takes time and future demand is uncertain, firms may increase their capacity and the volatility of output prices may increase with the number of firms.

These results are particularly salient in the context of an economy where competitive intensity is declining across a range of industries from 1980 through the present day[5]. Autor et al. (2017a) and De Loecker and Eeckhout (2017) document this transition; Galston and Hendrickson (2018) discuss potential policy responses. Several studies have provided evidence that this secular decline in competition is affecting macroeconomic dynamics. De Loecker and Eeckhout (2017) connect the increased market power to higher markups. Azar et al. (2017) and Benmelech et al. (2018) document the wage-setting market power of firms that dominate their local labour markets; as in our study, the authors focus on market power

---

[5] While Autor et al. (2017a) attribute much of the increase in competition to the dominance of “superstar” firms with high profits and a low labour share (Autor et al., 2017b). This explanation seems less applicable in the context of homebuilding as the industry is characterized by few productivity innovations that could drive changes to the labour share (McKinsey Global Institute, 2017). Section 1.1 discusses some likely explanations for the increase in market concentration in residential construction.

within local geographic markets. Gutiérrez and Philippon (2017) find a relationship between high market concentration and lower aggregate investment in industries in the United States in the 2000s. The present study focuses on a single industry. However, in light of the integral role of housing in consumption and investment decisions, changes in the production of housing have profound consequences for the broader economy. We examine the national impact of growing local market concentration in Section 4.2.

This study also brings a new perspective to a literature that seeks to understand the connection between housing construction and price dynamics. As reviewed in Gyourko (2009), recent improvements in data availability and the experience of the dramatic housing market cycle in the 2000s have motivated research on the determinants of housing production. Equilibrium models of the housing market that explicitly include the development sector include Poterba (1984), Topel and Rosen (1988), DiPasquale and Wheaton (1994), Grimes and Aitken (2010), and Liu et al. (2016). Recent research has also addressed the optimal development time. Among others, Mayer and Somerville (2000b) builds on theoretical results from Capozza and Helsley (1990) to estimate the parameters of a structural model and Murphy (2015) studies the firm's optimal timing and quantity construction decisions. Bulan et al. (2009) find that competition matters for the timing of development decisions by reducing the importance of volatility in determining when to build. Epple et al. (2010) and Combes et al. (2015) estimate housing production functions assuming price-taking atomistic firms. This study contributes to this literature by showing the integral role of market concentration in firms' production decisions.

Finally, this study provides insight into observed differences in the response of housing supply to changes in demand across different cities. Previous research attempts to explain cross-sectional differences in the supply of new housing in terms of scarcity of buildable land (Saiz, 2010) or regulatory constraints (Mayer and Somerville, 2000a; Green et al., 2005; Glaeser et al., 2008). Particularly relevant to the present research are Glaeser et al. (2008) and Paciorek (2013) which connect these supply constraints to the intensity of housing market cycles. The level of competition in the housing construction industry varies greatly across local housing markets. However, to our knowledge, the empirical literature has not previously examined the relationship between market concentration and cross-sectional variation in housing market cycles.

1.3 Novel contributions

In this study we use a novel panel data set to demonstrate the effect of market concentration on housing market dynamics. Specifically, we show that the secular decline in competitive intensity over the last decade has led to fewer housing units built, a decline in the inventory of vacant unsold housing units, and greater price volatility.

To provide economic context for our results we construct a theoretical model of oligopolistic firms choosing the timing of irreversible construction decisions in the context of uncertain future demand. This simple model is intended to capture stylized theoretical results in the real options literature (as introduced by Spatt and Sterbenz (1988), Grenadier (1996), Grenadier (2002), Aguerrevere (2003), and others) in a two-period model with multiple symmetric firms. As in those studies, the model predicts a greater rush to build early in markets with greater competitive intensity. Firms' rush to build leads to greater production volumes but lower price volatility.

We use a data set that identifies the firm that produces each new unit of housing sold in our sample area to document the rising market concentration in the residential construction market and investigate its impact on market dynamics. While this data set is widely used in the private sector, this study appears to be the first economic research to use this data set in academic research. We develop an instrument for the intensity of concentration to identify the causal relationship between market concentration and housing market cycles.

Our results indicate that more competitive markets have higher levels of production and a greater tendency towards overbuilding. This study is the first to empirically test this widely held qualitative intuition. However, we also show that more competitive markets have less drastic price fluctuations as greater competition reduces firms' ability to set prices appreciably above marginal cost.

To understand the economic magnitude of these results, we use our parameter estimates to investigate a counterfactual scenario where housing market competition remains at its high pre-recession level across the United States. Under this counterfactual, market outcomes would be very different. The annual volume of new housing would be $144 billion higher (equivalent to or 19% of net private fixed investment or 0.8% of gross domestic product). Approximately 681,000 additional housing units would be in the production pipeline. Housing price volatility would decline by over 50%.

The remainder of this article is organized as follows. First, we introduce a theoretical model that provides a framework for understanding how the rush to build can lead to overbuilding in a market with growing demand and high competition. Then, we introduce the data set and outline the empirical approach for testing the model's theoretical predictions. Finally, we present empirical results and describe the economic significance of our estimates in the context of the national housing market.

2 Theoretical model

This model illustrates the strategic behaviour of firms competing to provide housing in a market with an upward-sloping supply curve for land and a downward-sloping demand curve for housing. It provides a comparative static examination of how market outcomes vary with the number of active firms.

The model focuses on forward-looking firms' production decisions. Specifically, the environment consists of n symmetric firms producing housing over two periods. Firms make a single irreversible decision to build at either $t = 1$ or at $t = 2$. Firms use a Leontieff production technology to combine one unit of land and one unit of materials to produce one unit of housing. Each firm purchases land on a spot market subject to an upward-sloping supply curve at unit price ℓ, combine the land with materials at exogenous unit cost c to produce housing, and sell the housing at unit price p subject to a downward sloping demand curve. That is, conditional on choosing to build at t, the firm seeks to maximize $(p_t - c - \ell_t)\, h$ taking into account its own impact on p_t and ℓ_t.

Because housing is highly durable and land supply is fixed, both land and housing prices at $t = 2$ are affected by decisions at $t = 1$. Let H_t be the total volume of housing built by firms building at period t. Then, the land price ℓ_t and house price p_t are as follows for $t \in \{1, 2\}$:

$$\begin{aligned} \ell_1 &= \beta_0 + \beta_1 H_1 \\ p_1 &= \alpha_0 + \alpha_1 (Z - H_1) \end{aligned} \tag{1}$$

$$\begin{aligned} \ell_2 &= \beta_0 + \beta_1 (H_1 + H_2) \\ p_2 &= \alpha_0 + \alpha_1 (2Z - H_1 - H_2) \end{aligned} \tag{2}$$

In Equations 1 and 2, the price of housing includes an exogenous demand shifter. Its value grows over time from Z to $2Z$. Each firm takes into account its own impact on the supply curve for land and on the demand curve for housing. Accordingly, each firm faces a tradeoff between building at $t = 1$ and $t = 2$. At $t = 1$ land is more plentiful and housing stock which could compete with the firm's output has not yet been built. Conversely at $t = 2$ the demand for housing is higher.

While the model environment comprises only two periods, it captures the intuition of pre-emption and volume decisions by forward-looking homebuilding firms. At any point, firms are effectively in period $t = 1$ facing a given land supply and housing demand curve and deciding whether to build immediately or to wait for the realization of demand growth.

Before proceeding, it is helpful to introduce a normalization convention and some new notation that will clarify expressions later in the text. Specifically, we normalize $\alpha_1 Z \equiv 1$ and let $K = \alpha_0 + \beta_0 - c$. As well, the following parametric restriction will become necessary to ensure positive construction in equilibrium:

Assumption 1. *The supply curve for land, the cost of construction, and the demand curve for housing satisfy* $K > \frac{3}{2}$.

Qualitatively, this assumption ensures that the construction cost c is not so high relative to the cost of land and the price of housing that firms are unable to generate positive profits.

The solution concept in this model is a symmetric mixed-strategy weak perfect Bayesian equilibrium. The focal firm has beliefs regarding whether the other firms will build at $t = 1$ or at $t = 2$. Specifically, the focal firm believes that a number $m \in [0, n-1]$ of the other firms will build at $t = 1$ and the other $n - m - 1$ firms will build at $t = 2$[6] In equilibrium, these beliefs will be consistent with the other firms' actions. We seek a mixed strategy equilibrium; the firm will have a nonzero probability of building at $t = 1$ and a nonzero probability of building at $t = 2$.

For a given focal firm, let $\tilde{h}_1$ and $\tilde{h}_2$ be the number of units built of each of the other firms conditional on building at $t = 1$ and $t = 2$. Then, let $h_1^*\left(m; \tilde{h}_1, \tilde{h}_2\right)$ and $h_2^*\left(m; \tilde{h}_1, \tilde{h}_2\right)$ be the best responses of the focal firm. Taking first-order conditions and rearranging yields the following best responses conditional on building at $t = 1$ and at $t = 2$:

$$\begin{aligned} h_1^*\left(m; \tilde{h}_1, \tilde{h}_2\right) &= \frac{1}{2\left(\alpha_1 + \beta_1\right)} \left(K + 1 - \left(\alpha_1 + \beta_1\right) m \tilde{h}_1\right) \\ h_2^*\left(m; \tilde{h}_1, \tilde{h}_2\right) &= \frac{1}{2\left(\alpha_1 + \beta_1\right)} \left(K + 2 - \left(\alpha_1 + \beta_1\right) \left(m \tilde{h}_1 + \left(n - m - 1\right) \tilde{h}_2\right)\right) \end{aligned} \tag{3}$$

As we are interested in a mixed-strategy equilibrium, we seek a situation where firms are indifferent between building at $t = 1$ and at $t = 2$. Imposing symmetry on the decisions

---

[6] Throughout, we consider only the $n > 1$ case. If $n = 1$, in equilibrium the monopolist firm will always wait until the second period to build.

of the focal firm and the other firms and rearranging yields the following expression for the difference in optimal profits $\pi_1^*(m)$ and $\pi_2^*(m)$ between construction at $t = 1$ and $t = 2$ as a function of m:

$$\pi_2^*(m) - \pi_1^*(m) = \frac{-1}{4(n-m)^2(n+1)^2(m+1)} \left[m(-2-n) + 3n + 1 - K - 3n^2\right] \times$$
$$\left[m^2(K+5n+3) + m\left(-5Kn = 2K - 5n^2 + 3n + 2\right) + 2Kn^2 - Kn - K - 3n^2 + 2n + 1\right] \tag{4}$$

In a mixed-strategy equilibrium, the left-hand side of Equation 4 must be equal to zero. The term in square brackets on the first line of Equation 4 has no root with $m > 0$ under Assumption 1. However, the term in square brackets on the second line of Equation 4 has roots as follows:

$$m_\pm^* = \frac{1}{2(K+5n+3)} \left[5n^2 + (5K-3)n + 2(K-1) \pm\right.$$
$$\left.\sqrt{\left(5n^2 + (5K-3)n + 2(K-1)\right)^2 - 4(K+5n+3)\left(2Kn^2 - Kn - K - 3n^2 + 2n + 1\right)}\right] \tag{5}$$

It remains to show that Equation 5 describes a valid equilibrium belief — that is, a belief which is supported by a mixed-strategy equilibrium. The solution with the positive sign gives $m_+^* > n$ which is not a valid equilibrium belief. The following lemma will begin to establish that m_-^* does constitute a valid equilibrium belief.

Lemma 1. *The solution m_-^* to Equation 5 is positive.*

Proof. See Appendix A.1. □

The following lemma establishes the large-n behaviour of m_-^*:

Lemma 2. *As n grows large, $\frac{m_-^*}{n}$ is bounded above by $\frac{1}{2}$.*

Proof. See Appendix A.2. □

In the large-n limit, half the firms are building at $t = 1$ and half are building at $t = 2$. This fraction arises from the growth in the demand shifter from Z to $2Z$. Uneven growth would give a different limit but the intuition would remain unchanged. The following proposition establishes that m_-^* is monotonically increasing in n:

Proposition 1. *The equilibrium beliefs about the number of firms building in the first period* m^*_- *increases with* n *sufficiently quickly that* $\frac{m*_-}{n}$ *is increasing in* n*.*

Proof. See Appendix A.3. ☐

This proposition corresponds to a "rush" to build at $t = 1$. Although demand will be higher at $t = 2$ (and a monopolist would choose to build at $t = 2$), firms believe that their competitors will build at $t = 1$. If their competitors build at $t = 1$, the remaining land will be more expensive and the demand will be lower at $t = 2$. Accordingly, firms shift production to $t = 1$ with positive probability. In equilibrium, these beliefs are self-fulfilling. While the model represents a significant abstraction from reality, this result captures the real-world rush to purchase land, build housing, and capture market share.

From Proposition1, the following existence result follows directly:

Proposition 2. *For any number of firms* $n > 1$ *a mixed-strategy equilibrium characterized by* $m^*_- \in (0, n-1)$ *exists.*

Proof. See Appendix A.4. ☐

For the remainder of the discussion we will consider the equilibrium generated by belief m^*_-. For legibility we suppress the superscript and subscript and denote this belief by m. Imposing symmetry on Equation 3 yields the following construction decisions for each firm:

$$\begin{aligned} h_1^* &= \frac{K+1}{(m+2)(\alpha_1+\beta_1)} \\ h_2^* &= \frac{m+2K-4}{(m+2)(\alpha_1+\beta_1)(n-m+1)} \end{aligned} \tag{6}$$

This implies that the equilibrium aggregate production of housing in each period is as follows:

$$\begin{aligned} H_1^* &= \frac{mn}{n-1}\frac{K+1}{(m+2)(\alpha_1+\beta_1)} \\ H_2^* &= \frac{n^2-mn-n}{n-1}\frac{m+2K-4}{(m+2)(\alpha_1+\beta_1)(n-m+1)} \end{aligned} \tag{7}$$

From Equation 7 we can derive an additional theoretical result:

Proposition 3. *The total volume built at* $t = 1$ *is increasing in* n*.*

Proof. See Appendix A.5. ☐

According to Equation 6, the volume of construction by each firm at $t = 1$ is decreasing in n. However, according to Proposition 3, the rush to build shown in Proposition 1 is sufficiently large that increasing the number of firms increases the total volume of construction at $t = 1$. This result may seem unsurprising in light of Proposition 1 but it is worth emphasizing that this result would not arise in a marketplace of atomistic price-taking firms.

This discussion has focused on *ex ante* price and construction decisions. However, insofar as each firm is playing its mixed strategy independently, the *ex post* outcome varies with the realization of the n firms' mixed strategies. The following proposition demonstrates that an increase in the number of firms leads to a lower dispersion in *ex post* outcomes:

Proposition 4. *Assume m satisfies the restriction $\frac{2}{n^2+3} > \frac{m}{n(n-1)} > \frac{1}{n(n+1)}$. Then, the ex post price volatility at $t = 1$ is decreasing in n.*

Proof. See Appendix A.6. ☐

Remark 1. It is worth noting that the interval described by the two bounds in Proposition 4 is not empty. To see this, note that $\frac{2}{n^2+3} - \frac{1}{n^2+n} = \frac{(n+3)(n-1)}{n(n+1)(n^2+3)}$ which is positive for $n > 1$

Qualitatively, this proposition shows that the dispersion of prices decrease as more firms enter the market. The conditions in the proposition are sufficient, but not necessary. The upper bound on n excludes situations where the number of firms is so large that the market is close to the competitive limit and the dominant effect of an additional firm is the reduction in production by each firm. The lower bound on n excludes situations where the probability of any firm building at $t = 1$ is sufficiently low that the effective price is very close to $1 + \alpha_0$ and the volatility is very close to zero; any marginal increase in competition would raise the volatility.

3 Empirical approach

We use reduced-form instrumental variable regressions to understand the relationship between competitive intensity and market outcomes. Our outcome variables are closely related to Propositions 1, 3, and 4 — they measure the volume, volatility, and pricing of housing production. For each outcome variable y, we estimate the following specification across firms j, markets m and years t:

$$y_{mt} = \beta COMP_{mt} + \gamma X_{mt} + \lambda_m + \mu_t + \kappa_j + \nu_{mt} + \varepsilon_{mtj} \tag{8}$$

In Equation 8, $COMP_{mt}$ is a measure of competition intensity and X_{mt} is a vector of covariates including a measure of demand. All results include a full set of fixed effects:

1. Market fixed effects λ_m control for persistent differences in regulatory environment, existing land uses, and other time-invariant characteristics.
2. Year fixed effects μ_t control for macroeconomic conditions.
3. Firm fixed effects κ_j control for the possibility that different firms may have systematic differences in their construction costs, access to capital, or other variables influencing firm strategy.

When estimating Equation 8 we cluster standard errors at the market-year level as this is the level of variation in competition.

3.1 Market definition

We define markets for new housing following the Census definition of places. Places include incorporated cities, towns, and boroughs (e.g. Camden, NJ or Annapolis, MD) as well as Census-designated places in areas without formal municipal boundaries (e.g. Columbia, MD or Levittown, PA). To exclude very small markets which may have limited construction activity, we only include places with a 2015 population of at least $25,000$ (U.S. Census Bureau, 2015a). Places are a reasonable market delineation for new housing as they approximately match the spatial range over which consumers search for new housing.

Previous literature in housing markets has used the metropolitan statistical area (MSA) as a unit of analysis. However, this does not appear to be a plausible market delineation from the point of view of the homebuying consumer. As noted by Rozenfeld et al. (2011) and others, MSAs are composed of collections of counties and therefore reflect a combination of historical political boundaries and modern economic conditions. Moreover, counties are very large and heterogeneous in terms of income, built environment, travel time to work, and other attributes.

Zillow's zipcode-level indices of housing price per square foot provides an opportunity to test whether price dynamics within an MSA are reflective of a single integrated market (Zillow Research, 2018). We use this data to calculate the month-over-month percentage price changes from January 2006 onwards for all zipcodes in the largest MSA in our sample (Washington-Arlington-Alexandria, DC-VA-MD-WV). The median pairwise correlation between two zipcodes is 0.43 and over 2.5% of zipcode pairs are negatively correlated. However, for zipcode pairs in the same place as each other, the median correlation is 0.74 and less than 0.8% of pairs are negatively correlated. A similar comparison for the next-largest MSAs (Philadelphia-Camden-Wilmington, PA-NJ-DE-MD and Baltimore-Columbia-Towson, MD) yields a similar increase in correlation when only considering same-place zipcode pairs. Insofar as prices within a single market should co-move (in the classical economic sense, markets should obey the "law of one price"), this would seem to indicate that places reflect a more reasonable market delineation than MSAs.

Examples from our sample area illuminate the heterogeneity within MSAs. Median household income in the city of Baltimore is $42,241 while the median in Ellicott City (another place in the same MSA) is $114,916 (U.S. Census Bureau, 2015a). While counties are smaller than MSAs they are also often large and heterogenous. In Baltimore County (which had a Census population estimate of 831,128 as of 2015) Hampton had a 2015 median household income of $114,821 while Essex had a 2015 median household income of $48,434 (U.S. Census Bureau, 2015a). These large differences are highly salient to market delineation. As documented by Landvoigt et al. (2015), Liu et al. (2016) and others, the market dynamics for expensive and inexpensive housing within the same county can differ sharply. Moreover, counties often have shapes drawn for administrative convenience in previous centuries rather than natural communities; Baltimore County and Prince George's County wrap around the city of Baltimore and the District of Columbia and Montgomery County wraps around the northwest part of the city of Philadelphia.

Outside of the large urban centres in the sample, places also generally coincide more closely with school catchment boundaries. For example, in Maryland, Gaithersburg High School, Rockville High School, Owings Mills High School, Catonsville High School, and Glen Burnie High School all serve catchment areas which closely resemble the boundaries of their respective cities and Census-designated places. Households with children are sensitive to the quality of local public schools (Barrow, 2002) and therefore these catchment boundaries are salient to residential location decisions.

Moreover, using places as the definition of real estate markets is more in keeping with the empirical results on buyers' housing search preferences presented in Piazzesi et al. (2015). The authors find that one-quarter of prospective buyers in the Bay Area search in only a single zipcode. The remaining three-quarters tend to search among a tight cluster of zipcodes; the median search has a mean geographic distance of 3.2 miles and a car travel time of 13.1 minutes between zipcode centroids. This is comparable in scale to the places in our sample.

Under this definition, the majority of our markets are suburban communities. Despite the highly visible construction in dense urban centres, new housing remains concentrated in the suburbs. In the second half of the twentieth century nearly all growth in urban population and land area occurred in the suburbs (Nechyba and Walsh, 2004). Between 2000 and 2010, the population grew faster in the suburbs than in the city centre in 98 of the hundred largest metropolitan areas (Couture and Handbury, 2017) and Census data suggests that this pattern has continued in large metropolitan areas since 2010 (Frey, 2018). Accordingly, our largely suburban market definition reflects the real-world geography of housing construction.

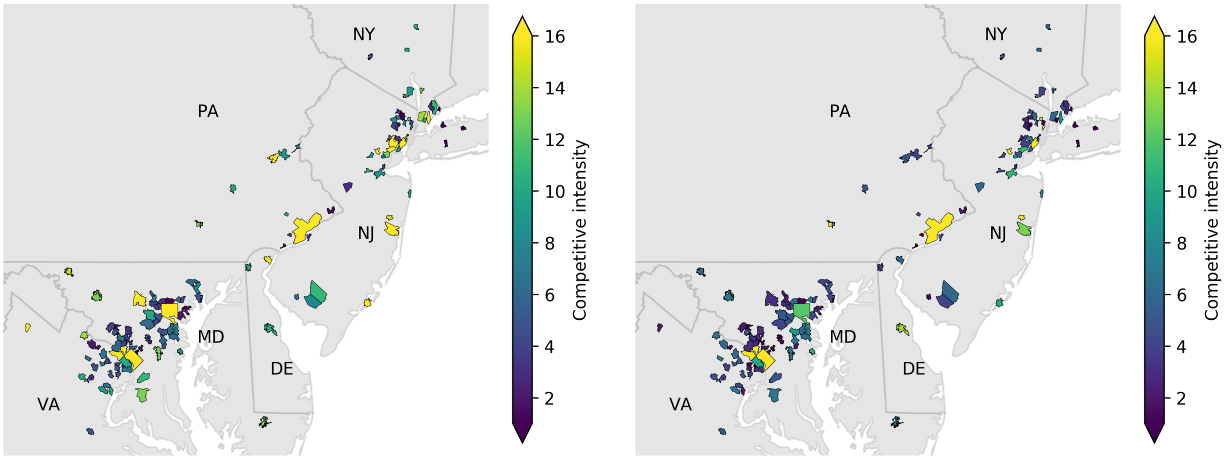
3.2 Measuring competition

We use a data set on residential construction that appears to be novel to the academic literature to construct a measure of competitive intensity. We develop an instrumentation strategy to account for potential endogeneity between housing market activity and residential construction.

3.2.1 Metrostudy data

To quantify market competition and understand how firms respond to their market power and the market power of their competitors, we require a data set that identifies individual firms' production decisions. Metrostudy (a proprietary data service operated by the firm Hanley Wood) collects exactly this data by monitoring residential development and property transactions to identify the firms building and selling individual housing units at a fine level of spatial disaggregation (Metrostudy, 2018). Their data collection covers all housing sales in the past eleven years across a broad geographical extent including all large metropolitan areas and approximately three-quarters of the United States population. This unique data set contains information on individual firms' construction activity in all markets (including the price, quantity, and attributes of new housing units) that is not available from other sources. Metrostudy also includes information on firms' characteristics including the overall scale of their operations. While this data set is widely used in private industry, it appears that it has not been used in the academic literature to date.

Our sample consists of places in the metropolitan areas centered on Philadelphia, Washington DC, Baltimore, Allentown, PA, Atlantic City, NJ, Dover, DE Salisbury, MD, Trenton, NJ, and Vineland, NJ as well as large parts of the New York metropolitan area excluding New York City itself. Figure 4 displays the markets included in the sample. As mentioned



(a) 2006

(b) 2015

Figure 4: Markets included in the Metrostudy data. The degree of competition shown on this map is the number of firms accounting for 90% of construction.

above we exclude places with less than $25,000$ residents. This leaves 132 markets with a total population of 9.74 million (U.S. Census Bureau, 2015a).

The theoretical model described in Section 2 connects the number of firms active in the market to various market outcomes. To construct an equivalent measure in the data we consider the number of firms building 90% of new housing in a given year. As this excludes small firms building a very small number of highly customized luxury housing units, this effectively captures the number of firms active in the homebuilding market. Figures 1 and 4 show this measure for the markets in our sample.

3.2.2 Instrumenting for concentration

We are measuring the effect of concentration on housing market outcomes including the timing of construction, price, and quantity of new housing. However, it is possible that that market concentration is an endogenous regressor in Equation 8 — that is, that $\operatorname{cov}(COMP_{mt}, \varepsilon_{mtj}) \neq 0$. This endogeneity could arise from local shocks that simultaneously impact competition and housing market outcomes. For example, a change to permit approval policy could affect both competition (through changes to barriers to entry) and the number of units produced. To mitigate this concern, we construct an instrumental variable for the level of competition.

In order to provide causal identification, the variation in the instrument must be driven by variation exogenous to local market conditions. The activity of firms in our data is affected by both local and national considerations. Our instrument is based on the local impact of variation in activity resulting from events at the national level. As discussed in Section 1.1 large national firms have been affected over our sample period by a set of macro-level circumstances which have favoured large national firms and enabled them to increase their local market power.

Specifically, we construct an instrument from the predicted behaviour of national firms[7]. We forecast the construction activity of a national firm j in market m using the activity of firm j in all markets other than m. This instrumentation strategy relies on the fact that these national firms have residential construction activity distributed across the United States and are particularly influenced by the changes to the national environment outlined in Sections 1.1. This strategy follows the widely-used shift-share instrumentation strategy introduced to the economic literature by Bartik (1991) and Blanchard et al. (1992). Moreover, the rationale for this instrument is similar to the rationale for the instrument for competitive intensity introduced in Atalay et al. (2017): from the point of view of a very large nationally active firm, individual cities are effectively negligible.

To explicitly define this instrument, let C_{mtj} be activity of firm j at period t in market m. Also, let J^N_{mt} be the set of national firms active at period t in market m (i.e., the set of national firms with nonzero closings in this market and this year). Then, define $\hat{C}_{mtj}$ as the predicted activity by firm j at period t in market m where the prediction comes from the activities of firm j in all markets other than m:

$$\hat{C}_{mtj} = \frac{\sum_{m' \neq m} C_{m'tj}}{\sum_{m' \neq m} C_{m',t-1,j}} C_{m,t-1,j} \tag{9}$$

Next, for each market, define the market-weighted average over all national firms $\bar{C}_{mt}$:

$$\bar{C}_{mt} = \frac{\sum_{j \in J^N_{mt}} C_{mt,j-1} \hat{C}_{mtj}}{\sum_{j \in J^N_{mt}} C_{mt,j-1}} \tag{10}$$

Finally, to obtain an instrument for the number of firms accounting for 90% of production, normalize by the previous year's construction:

$$Z_{mt} = \frac{\bar{C}_{mt}}{\sum_{j \in J^N_{mt}} C_{mt,j-1}} \tag{11}$$

We use Z_{mt} as an instrument for competitive intensity in the regression results presented

---

[7] The Metrostudy data set categorizes each firm as national, regional, local, or micro depending on their total production. The data set includes 42 nationally-active firms.

below. One would expect negative correlation between the instrument and the observed competitive intensity $COMP_{mt}$. For example, if one large national firm increased its production sharply, then Z_{mt} would rise while the number of firms accounting for 90% of production would fall. This intuition bears out in reality; the unconditional correlation between the instrument and $COMP_{mt}$ is negative (-0.094) and statistically significant. When reporting results in Section 4.1 we include the first-stage F-statistic throughout. Because our model includes only one endogenous regressor, this is equivalent to the Cragg-Donald statistic described in Cragg and Donald (1993). In every regression reported in this study, we reject the possibility of a weak instrument according to the tests described in Stock and Yogo (2005). We also report the p-value of the coefficient on the instrument in the first stage of the regression.

Our instrumentation strategy focuses solely on the activity of large national firms. It is worth emphasizing that these firms are present and highly active across the markets in our sample. Figure 5 shows the cumulative distribution of the share of production accounted for by national firms in the market-year pairs in our sample. As shown, national firms produce housing in 74% of market-year pairs in the sample. They account for at least 10% of production (and therefore impact the number of firms accounting for 90% of production) in 54% of market-year pairs and account for the majority of construction in 15% of market-year pairs.

3.3 Measuring demand

To estimate Equation 8, we require a measure of demand for housing in market m at time t. Ideally, this measure should be exogenous to the condition of the housing market in m at t. We use the number of jobs accessible from market m as a measure of demand. Specifically, we calculate the number of jobs within fifty miles[8] of housing market m in the Quarterly Census of Employment and Wages (Bureau of Labor Statistics, 2017). To avoid potential endogeneity between economic activity in market m and housing construction in market m we follow papers including Bayer et al. (2007) by only considering jobs outside the county in which m is located.

It is worth emphasizing that we are not using this exogenous demand shifter as an

---

[8] While this may seem like a lengthy commute, it is not unreasonable in the context of the housing markets in our sample. For instance, 5% of workers in Fredericksburg City, VA, 11% of workers in Orange County, NJ, and 12% of workers in Suffolk County, NY travel to a county more than fifty miles away. Somewhat shorter commutes are even more common; the majority of workers in Loudoun County, VA commute to a county at least 25 miles away (McKenzie, 2013).

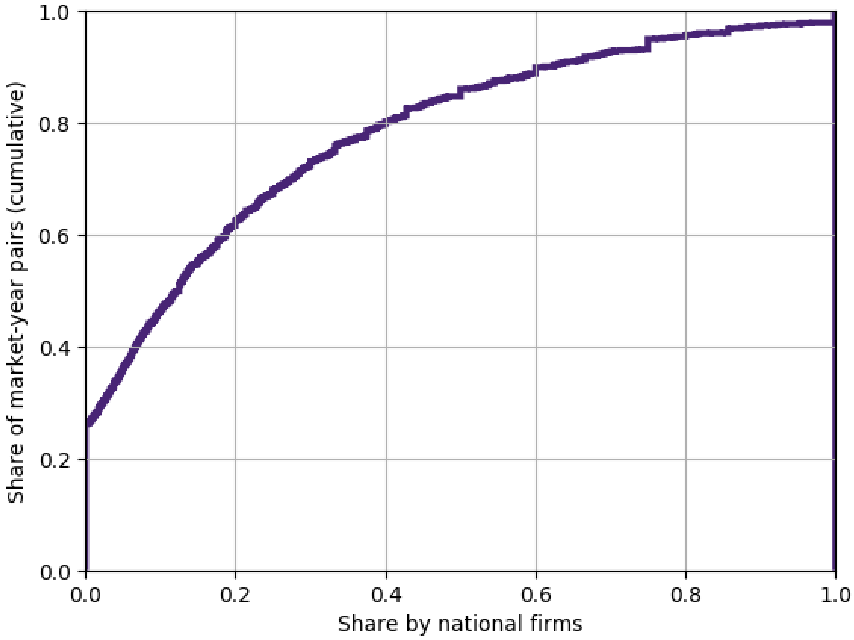


Figure 5: Cumulative distribution of the share of units accounted for by national firms across all market-year pairs.s

instrument for the underlying demand for housing. The underlying demand for housing is unobservable. Realized population growth reflects the equilibrium outcome of changing supply in response to demand. To avoid conflating demand growth with changes in supply, we include this plausibly exogenous demand measure directly as a regressor.

3.4 Measuring construction cost

We use data from RSMeans (2016) (as used by Gyourko and Saiz (2006) and others) to account for differences in the cost of construction across markets[9]. We use the "overall" index which comprises the total cost of construction including both materials and labour. The data set includes a price index for each three-digit zipcode and a price index for each year. We map these three-digit zipcodes to the markets in our sample and multiply the market index by the year index to obtain a value for each market-year pair.

---

[9] As the markets for the materials and labour needed for housing construction are larger than any individual place we regard the construction cost as exogenous.

3.5 Additional covariates

To understand how other differences in market composition could affect housing market dynamics we explore additional variables. As shown below, all regression results are robust to the inclusion of these additional variables.

To control for the possibility that firms operating at different scales interact differently, we control directly for the share of construction by national, regional, and micro-sized home building firms[10].

We also include a measure of "established" markets to control for markets where most new housing is produced through redevelopment. The dynamics of the housing development process may be systematically different in these markets. Specifically, we include an indicator variable for whether the market's resale share of total sales in a given year falls into the top tercile of all markets in that year.

4 Results and discussion

We present empirical results that demonstrate the impact of market concentration in homebuilding on housing market cycles. Then, we conduct a counterfactual exercise to quantify the impact of these changes on the scale of the macroeconomy.

4.1 Empirical results

We investigate the relationship between market concentration and cycle dynamics by estimating models of the form specified by Equation 8. We measure concentration using the number of developers accounting for 90% of construction in the market at year t; a higher value indicates greater competitive intensity. We show OLS coefficient estimates as well as the estimates with the instrumental variable for competitive intensity discussed in Section 3.2.2. Section 4.2 quantifies the economic significance of these results by investigating how changes in market concentration between 2006 and 2015 changed market outcomes. Appendix B shows results that demonstrate the robustness of our estimates to changes in sample and specification. Table 1 shows summary statistics for the data set used in these regressions.

---

[10] As discussed in further detail below, the Metrostudy data contains information on the overall scale of firms' operations.

Table 1: Summary statistics for the regression data. The unit of observation in this table is a market-year-firm tuple.

Variable	N	Median	Mean	Std. dev	Min	Max
Number of firms producing 90%	22344	12.0	25.08	31.08	1	129
Jobs within fifty miles (millions)	22344	3.54	4.33	2.16	0.15	8.49
Construction cost index	19096	1.94	1.96	0.25	1.38	2.52
National firm share	22344	0.08	0.15	0.18	0.00	1.00
Regional firm share	22344	0.08	0.12	0.13	0.00	1.00
Micro firm share	22344	0.67	0.61	0.25	0.00	1.00
Established markets	12419	0.00	0.25	0.43	0.00	1.00
Total value of housing ($ million)	22343	81.20	183.92	233.10	0.10	1120.00
Production volatility	22313	43.00	108.27	167.55	0.00	998.00
Housing units in pipeline (months)	12858	10.49	23.09	38.74	0.0	667.2
Price volatility (per square foot)	22170	15.93	30.61	54.97	0.08	1364.88
Firms per market-year	1437	10.0	15.55	22.80	1	225
National firms per market-year	1437	2.0	2.32	2.52	0	17
Markets	137					
Years	11					
Firms	8045					

Tables 2 and 3 show regression coefficients where the dependent variables are the logarithm of the total value of new housing built and the logarithm of the total square feet of housing built. As shown, markets with a greater degree of concentration produce significantly less housing even with after instrumenting for market concentration. This result is consistent with the prediction of Proposition 3 as well as the theoretical prediction in a static oligopolistic market. However, this result is difficult to reconcile with models that assume a competitive sector of atomistic price-taking construction firms.

Proposition 1 predicts a "rush" to build housing in markets with greater competitive intensity. We test for the rush to build using two outcome variables. Table 4 shows regression coefficients where the dependent variable is the logarithm of the year-over-year absolute change in the number of housing units produced while Table 5 shows regression coefficients where the dependent variable is the sum of the number of vacant units, model units[11], and units under construction — that is, the total housing supply in the pipeline. The latter measure is scaled to the size of the market by expressing the inventory in terms of months of supply at contemporaneous local absorption rates. As shown, the theoretical prediction is borne out; a greater degree of competitive intensity is associated with more drastic supply

---

[11] Model units are vacant unsold units used for display purposes for prospective buyers.

Table 2: Regression results where the dependent variable is the logarithm of the total value of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.064	1.10	0.067	1.36	0.064	1.10
	(0.013)	(0.18)	(0.014)	(0.26)	(0.013)	(0.18)
Jobs within 50 miles	0.79	7.52	0.48	7.22	0.79	7.52
	(0.50)	(1.33)	(0.50)	(1.49)	(0.50)	(1.33)
Construction cost	-0.33	-0.47	-0.26	-0.36	-0.33	-0.47
	(0.034)	(0.049)	(0.034)	(0.050)	(0.034)	(0.049)
Share national firms			1.06	0.72		
			(0.085)	(0.13)		
Share regional firms			0.23	0.23		
			(0.094)	(0.13)		
Share micro firms			0.41	-1.08		
			(0.076)	(0.31)		
Established market					-0.47	
					(0.56)	
Observations	15782	15782	15782	15782	15782	15782
R^2	0.776	0.640	0.781	0.594	0.776	0.640
1^{st} Stage F-statistic		87.567		55.303		87.567
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table 3: Regression results where the dependent variable is the logarithm of the total square footage of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.062	1.05	0.064	1.28	0.062	1.05
	(0.013)	(0.17)	(0.013)	(0.24)	(0.013)	(0.17)
Jobs within 50 miles	0.91	7.25	0.60	6.89	0.91	7.25
	(0.48)	(1.26)	(0.48)	(1.40)	(0.48)	(1.26)
Construction cost	-0.34	-0.47	-0.27	-0.36	-0.34	-0.47
	(0.032)	(0.046)	(0.032)	(0.047)	(0.032)	(0.046)
Share national firms			1.04	0.71		
			(0.081)	(0.13)		
Share regional firms			0.18	0.14		
			(0.090)	(0.12)		
Share micro firms			0.39	-1.01		
			(0.073)	(0.30)		
Established market					-1.39	
					(1.43)	
Observations	15768	15768	15768	15768	15768	15768
R^2	0.761	0.617	0.767	0.572	0.761	0.617
1^{st} Stage F-statistic		87.010		55.169		87.010
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

fluctuations as well as a larger supply of housing in the pipeline to construction.

Table 4: Regression results where the dependent variable is the logarithm of the absolute value of the change in the number of the units of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.18	1.46	0.20	1.74	0.18	1.46
	(0.034)	(0.38)	(0.035)	(0.52)	(0.034)	(0.38)
Jobs within 50 miles	-1.52	6.79	-2.42	5.69	-1.52	6.79
	(1.27)	(2.81)	(1.26)	(3.04)	(1.27)	(2.81)
Construction cost	-1.04	-1.21	-0.87	-1.00	-1.04	-1.21
	(0.085)	(0.10)	(0.086)	(0.10)	(0.085)	(0.10)
Share national firms			1.90	1.41		
			(0.22)	(0.29)		
Share regional firms			-0.18	-0.25		
			(0.24)	(0.26)		
Share micro firms			0.40	-1.44		
			(0.19)	(0.65)		
Established market					-3.90	
					(1.41)	
Observations	15703	15703	15703	15703	15703	15703
R^2	0.322	0.223	0.334	0.205	0.322	0.223
1^{st} Stage F-statistic		88.280		54.375		88.280
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Finally, Proposition 4 predicts that (under certain conditions) price volatility is decreasing in the level of competitive intensity. To test this, we use the absolute proportional change in housing price per square foot $\|p_t - p_{t-1}\| / p_{t-1}$ as the dependent variable. Table 6 shows the results of this regression. As predicted, higher competitive intensity is associated with substantially lower price volatility.

We conduct several tests of robustness to ensure the validity of these empirical results:

1. To ensure that our results are not driven by markets which may be smaller than the geographic extent of a typical homebuyer's search, we repeat the analysis excluding

Table 5: Regression results where the dependent variable is the months of supply of housing (including under construction, model and finished vacant) at contemporaraneous absorption rates.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.18	2.01	0.18	3.16	0.18	2.01
	(0.048)	(0.50)	(0.051)	(0.94)	(0.048)	(0.50)
Jobs within 50 miles	0.089	-6.83	0.42	-10.9	0.089	-6.83
	(3.82)	(5.23)	(3.82)	(7.06)	(3.82)	(5.23)
Construction cost	-7.12	-0.47	-6.90	2.73	-7.12	-0.47
	(0.61)	(1.95)	(0.62)	(3.19)	(0.61)	(1.95)
Share national firms			0.21	1.62		
			(0.29)	(0.64)		
Share regional firms			-0.47	1.75		
			(0.30)	(0.85)		
Share micro firms			-0.11	-1.45		
			(0.24)	(0.57)		
Established market					-17.6	
					(3.30)	
Observations	4379	4379	4379	4379	4379	4379
R^2	0.405	-0.015	0.407	-0.574	0.405	-0.015
1^{st} Stage F-statistic		32.015		15.499		32.015
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table 6: Regression results where the dependent variable is the logarithm of the absolute value of the change in price per square foot.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	-0.11	-3.15	-0.033	-3.78	-0.11	-3.15
	(0.036)	(0.50)	(0.038)	(0.72)	(0.036)	(0.50)
Jobs within 50 miles	-8.69	-28.2	-8.58	-28.0	-8.69	-28.2
	(1.36)	(3.68)	(1.36)	(4.15)	(1.36)	(3.68)
Construction cost	-0.23	0.16	-0.33	-0.049	-0.23	0.16
	(0.092)	(0.14)	(0.092)	(0.14)	(0.092)	(0.14)
Share national firms			-2.06	-1.05		
			(0.23)	(0.38)		
Share regional firms			-0.44	-0.32		
			(0.26)	(0.36)		
Share micro firms			-1.59	2.75		
			(0.21)	(0.88)		
Established market					-11.6	
					(1.48)	
Observations	15763	15763	15763	15763	15763	15763
R^2	0.180	-0.404	0.191	-0.600	0.180	-0.404
1^{st} Stage F-statistic		86.989		55.012		86.989
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

any market with fewer than forty thousand residents (U.S. Census Bureau, 2015a). These excluded markets comprise 30.1% of all observations. Appendix B.1 contains regression results generated with this reduced sample.

2. To ensure that our results are not driven by apparent high concentration in markets with very low volume, we repeat the analysis excluding all market-year pairs with production in the lowest decile of nonzero housing production across all markets in each year. These excluded markets comprise 9.3% of all observations. Appendix B.2 contains regression results generated with this reduced sample.
3. As defined in Equation 9, the instrumental variable uses the predicted activity of national firms in market m based on their activity in all markets $-m \neq m$. To account for the possibility that firms' activity is correlated across nearby markets, we also construct the instrument excluding not only the focal market m but also all other markets spatially adjacent to m. Appendix B.3 contains regression results generated with this alternate instrument.

Under each of these alternative specifications, the magnitude and significance of our empirical results remain qualitatively similar.

As an additional check on the power of our instrumental variable in a context with many fixed effects we randomly reassign the values of the instrument across observations 400 times. Under this random shuffling, our results should no longer hold. Across all outcome variables this reshuffled placebo instrument generates significance at $p < 0.05$ only 7.8% of the time. This provides additional confidence that our instrument has adequate power.

Taken together these empirical results support the intuition widely held by the private sector that the decline in competitive intensity in local housing construction has altered the dynamics of the real estate cycle. More concentrated markets feature lower levels of production, less volatility in production, and greater volatility in prices. These effects are statistically significant and economically meaningful.

4.2 Counterfactual exercise

In order to understand the macroeconomic significance of our results, we consider a counterfactual scenario where the level of competition in the housing market in 2015 remained at its higher 2006 level. Suppose that, absent the developments discussed in Section 1.1, the level of competition had held constant at pre-recession levels in markets throughout the United States. What would this imply for the state of the present housing construction market?

As discussed in Section 3.2.1, we have information on local competitive intensity and market outcomes in only a subsample of the country. We adopt the following process to infer competitive intensity for the rest of the country using the Zipcode Business Patterns data set (U.S. Census Bureau, 2015b) which provides information on zipcode-level employment in residential construction for 2012 through 2015:

1. We aggregate Zipcode Business Patterns data to the markets in our sample using GIS software.
2. We generate a measure of the implied concentration in the Zipcode Business Patterns data by assuming that production increases linearly with the number of employees.
3. For the subsample of years and markets for which we have Metrostudy data we estimate a mapping from the implied concentration from Zipcode Business Patterns to the measured concentration from Metrostudy using a flexible polynomial specification.
4. We use the mapping generated in Steps 2 and 3 to predict the level of competitive intensity in 2015 for all markets in the United States.
5. For the markets in the Metrostudy data, we estimate a mapping between 2006 and 2015 level of concentration using a flexible polynomial specification.
6. We use the result of Step 4 and mapping generated in Step 5 to predict the level of competitive intensity in 2006 for all markets in the United States.
7. For each outcome variable we use the coefficients in the second column of the tables presented above to estimate the impact of changes in competitive intensity under a counterfactual scenario market competitive intensity in 2015 held at 2006 levels.

Figure 6 compares the distribution of competitive intensity (as inferred from Zipcode Business Patterns data) for the markets in our sample with the distribution for the entire United States. As shown, the fit from Zipcode Business Patterns data suggests that the distribution of competitive intensity in our Metrostudy sample is similar to the distribution for the overall United States. This correlation is robust to the choice of assumption used to map from Zipcode Business Patterns to market concentration described in Step 2.

However, while the distribution of competitive intensity in our sample is similar to the distribution in the United States overall, the markets in our sample may differ from other markets in other ways. For example the markets in our study appear to be in the middle

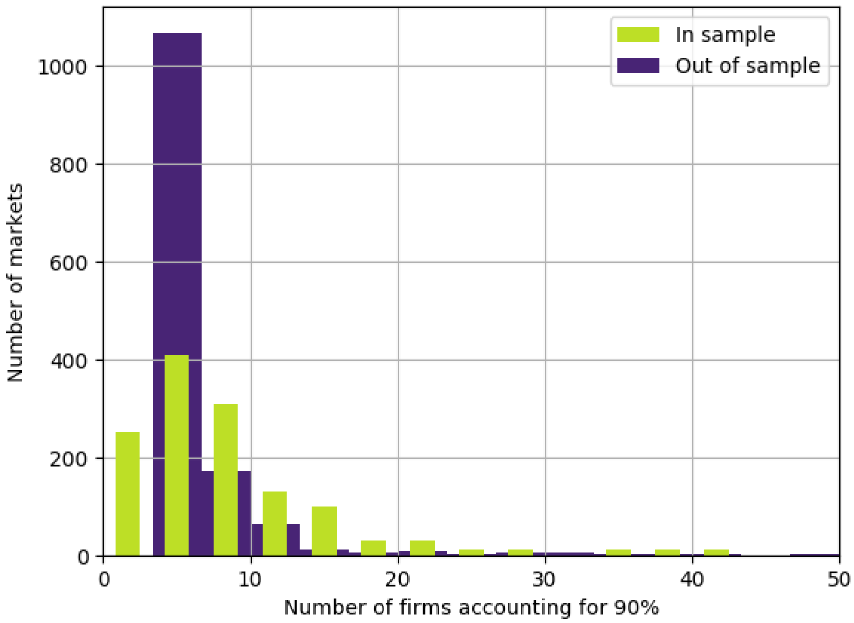


Figure 6: Predicted number of firms accounting for 90% of housing construction in each market in the United States based on 2006. Predicted values obtained from fit to Zipcode Business Patterns data. Narrow green bars show Metrostudy markets and broad blue bars show the rest of the United States. The vertical scale of the narrow green bars is increased by a factor of ten for visual clarity.

of the national distribution for housing supply elasticity[12] and therefore these markets may not experience identical dynamics to highly constrained markets or very unconstrained markets. Moreover, our analysis focuses on markets with at least 25,000 residents and therefore excludes small towns and rural areas. We acknowledge these potential limitations to the external validity of this counterfactual exercise.

Under this set of assumptions, the median number of firms accounting for 90% of production across all markets in the United States from 2006 through 2015 fell from 6.25 to 4.78 — i.e., a decrease of 24%. Weighting the markets by pre-period population (as measured in the 2015 five-year American Community Survey estimates) does not appreciably alter this result. Table 7 shows the predicted impact on markets across the United States at the quartiles of the distribution of predicted levels of 2006 competition. All changes in outcome variables are relative to the predicted levels of 2015 competition. As shown, the impact is

---

[12] Saiz (2010) ranks metropolitan areas according to their housing supply elasticity. Seven of these metropolitan areas have principal cities included in our Metrostudy sample. The supply elasticity rankings for these seven metropolitan areas range from the 27th percentile to the 60th percentile of the national distribution.

Table 7: Predicted competition levels and corresponding changes in value of housing supply and months of finished housing supply in all markets in the United States, evaluated at the 25th percentile, median, and 75th percentile of predicted levels of 2006 competition.

	25th	Median	75th
Concentration in 2006	6.22	6.25	6.44
Concentration in 2015	4.63	4.78	5.47
Δ Value of housing supply (%)	38	34	26
Δ Square footage produced (%)	54	48	36
Δ Supply in pipeline (%)	81	71	53
Δ Price volatility (%)	-49	-57	-60

relatively uniform across the distribution of competitive intensity.

These estimates imply that the decrease in competition has impacted housing markets in economically meaningful ways. The total value of private residential construction in 2015 was $423 billion (U.S. Census Bureau, 2017b). Estimates from Table 7 indicate that absent the decrease in competition, the total value of housing would be on the order of $144 billion greater. The $144 billion difference in construction value is equivalent to 19% of the value of net private fixed investment in the United States economy in 2015 (Bureau of Economic Analysis, 2017a). Moreover, this is equivalent to 0.8% of 2015 GDP. This is comparable to the scale of the decline in residential investment Leamer (2007) identifies prior to previous recessions.

In 2015, 529,000 new single-family units intended for sale started construction and 398,000 new multi-family units intended for sale or rent started construction for a total of 857,000 new units under construction (U.S. Census Bureau, 2017a). Also, in 2015 12% of all housing units in the United States were vacant U.S. Census Bureau (2018) which yields an estimate of 960,000 housing units under construction or vacant and unsold. Estimates from Table 7 suggest that under this counterfactual level of competitive intensity the number of units in the supply pipeline would increase by approximately 681,000 units.

Our results also suggest that price volatility would be much lower under the counterfactual 2006 levels of market concentration. Data from Zillow suggests that between 2013 and 2017 the average absolute annual change in house prices at the market level was 5.5%. According to our estimates, this would be reduced to approximately 2.4% under 2006 competitive intensity.

It is worth emphasizing that these comparative static calculations likely understate the macroeconomic impact of concentrated housing construction markets and its effects on housing supply. As Hsieh and Moretti (2015), Ganong and Shoag (2017), and Bunten (2017) have noted, local constraints on housing supply can drive misallocation of labour and this misallocation can have substantial macroeconomic consequences.

5 Conclusion

We use a novel data set to examine the relationship between market concentration and market volatility in the residential construction sector. This data set allows us to document for the first time the high and rising market concentration at the level of individual housing markets. Our empirical results indicate that a higher degree of concentration in local housing construction markets leads to less housing production, a decreased rush to build more units, and greater volatility in prices. These findings are compatible not only with the theoretical model in this study but also with the stylized results of the literatures on real estate cycles and oligopoly as well as private industry's understanding of real estate market cycles. Our parameter estimates imply that the increasing concentration in the production in housing has led to a substantial reduction in the volume of housing produced as well as in the inventory of new vacant units.

Our counterfactual exercise suggests that the increase in market concentration from 2006 through 2015 led to the production of $144 billion less housing per year. This equivalent to 19% of the value of net private fixed investment in 2015. The reduction in housing construction from the increase in local market concentration has meaningful effects on overall macroeconomic investment.

The empirical results of this study indicate potential future directions for macroprudential policy. Regulators in Hong Kong and Korea have attempted to stem housing speculation by setting loan-to-value limits that reflect the perceived degree of risk in residential mortgages (Lim et al., 2011); these rules appear to have had a meaningful impact on house price dynamics in Korea (Igan and Kang, 2011). The efficacy of these policy interventions is predicated upon policymakers' ability to identify the potential for price volatility in different markets. Our research indicates that this may be a particularly significant concern in markets with high levels of concentration.

The study also has implications for policy at the local level. Municipal and regional governments have implemented a wide range of strategies to increase the supply or lower the cost of housing but to date these policies do not appear to take into account the role of competition between builders in providing new housing (Kingsley and Williams, 2007;

Bellisario et al., 2016; Kalugina, 2016; MacDonald, 2016). Insofar as local governments can control the level of competitive intensity through permit allocation, our results indicate an novel channel for influencing the supply of new housing.

Housing market cycles are a central component of macroeconomic cycles. In this study, we demonstrate empirically for the first time the impact of local housing market concentration on housing market cycle dynamics. This research provides a direction for new empirical investigation into the housing-driven component of the macroeconomic cycle.

References

Aguerrevere, Felipe L, "Equilibrium investment strategies and output price behavior: A real-options approach," *The Review of Financial Studies*, 2003, *16* (4), 1239–1272.

Atalay, Enghin, Ali Hortaçsu, Mary Jialin Li, and Chad Syverson, "How Wide Is the Firm Border?," Technical Report, National Bureau of Economic Research 2017.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen, "Concentrating on the Fall of the Labor Share," *American Economic Review*, may 2017, *107* (5), 180–185.

_ , _ , Lawrence F Katz, Christina Patterson, John Van Reenen et al., "The Fall of the Labor Share and the Rise of Superstar Firms," Technical Report, Centre for Economic Performance, LSE 2017.

Azar, José, Ioana Marinescu, and Marshall I Steinbaum, "Labor Market Concentration," Technical Report, National Bureau of Economic Research 2017.

Barr, Colin, "Handsome government handout for homebuilders," *Fortune*, January 7 2010.

Barrow, Lisa, "School choice through relocation: evidence from the Washington, DC area," *Journal of Public Economics*, 2002, *86* (2), 155–189.

Bartik, Timothy, *Who Benefits From State and Local Economic Development Policies*, Kalamazoo, Mich.: W.E. Upjohn Institute for Employment Research, 1991.

Bayer, P., F. Ferreira, and R. McMillan, "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 2007, *115 (4)*, 588–638.

Beebe, Paul, "Back from brink, Utah's Woodside Homes downsizes to fit new realities," *Salt Lake Tribune*, December 24 2012.

Bellisario, Jeff, Micah Weinberg, Camila Mena, and Lanwei Yang, *Solving the Housing Affordability Crisis*, Bay Area Council Economic Institute, 2016.

Benmelech, Efraim, Nittai Bergman, and Hyunseob Kim, "Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages?," Technical Report, National Bureau of Economic Research 2018.

Blanchard, Olivier Jean, Lawrence F Katz, Robert E Hall, and Barry Eichengreen, "Regional evolutions," *Brookings papers on economic activity*, 1992, *1992* (1), 1–75.

Board of Governors of the Federal Reserve System, "Survey of Consumer Finances," 2017.

Bray, Chad and Matthew Goldstein, "Lennar Makes Deal for CalAtlantic as Home Builders Face Challenges," *New York Times*, October 30 2017.

Builder Magazine M&A, "Lennar Corp., CalAtlantic Group Agree to Merge," *Builder*, October 30 2017.

Bulan, Laarni, Christopher Mayer, and C Tsuriel Somerville, "Irreversible investment, real options, and competition: Evidence from real estate development," *Journal of Urban Economics*, 2009, *65* (3), 237–251.

Bunten, Devin, "Is the Rent Too High? Aggregate Implications of Local Land-Use Regulation," Technical Report, Federal Reserve Board working paper series 2017.

Bureau of Economic Analysis, "Gross and Net Domestic Investment by Major Type," 2017.

_ , "Personal Consumption Expenditures by Function," 2017.

Bureau of Labor Statistics, "Quarterly Census of Employment and Wages," 2017.

Capozza, Dennis R and Robert W Helsley, "The stochastic city," *Journal of urban Economics*, 1990, *28* (2), 187–203.

Clifford, Catherine, "Pulte Homes in $3.1 billion merger," *CNN Money*, April 8 2009.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, "The Production Function for Housing: Evidence from France," December 2015.

Cook, Christopher D., "Homebuilder Lennar uses federal taxpayer funds to balance its books," *San Francisco Public Press*, July 6 2010.

Corkery, Michael and Jesse Drucker, "Corporate-Tax Break Revives in Budget Plan," *Wall Street Journal*, March 6 2009.

Couture, Victor and Jessie Handbury, "Urban revival in America, 2000 to 2010," Technical Report, National Bureau of Economic Research 2017.

Cragg, John G and Stephen G Donald, "Testing identifiability and specification in instrumental variable models," *Econometric Theory*, 1993, *9* (2), 222–240.

Crocker, Robb, "Elliott Building Group Selling All Assets," *Builder*, October 23 2007.

Davila, Eduardo and Anton Korinek, "Pecuniary Externalities in Economies with Financial Frictions," *The Review of Economic Studies*, 2016.

DiPasquale, Denise and William C Wheaton, "Housing market dynamics and the future of housing prices," *Journal of urban economics*, 1994, *35* (1), 1–27.

Epple, Dennis, Brett Gordon, and Holger Sieg, "A new approach to estimating the production function for housing," *The American Economic Review*, 2010, *100* (3), 905–924.

Frey, William H., "US population disperses to suburbs, exurbs, rural areas, and "middle of the country" metros," Technical Report, Brookings Institute March 26 2018.

Galston, William A. and Clara Hendrickson, "A policy at peace with itself: Antitrust remedies for our concentrated, uncompetitive economy," Technical Report, Brookings Institute January 5 2018.

Ganong, Peter and Daniel Shoag, "Why has regional income convergence in the US declined?," *Journal of Urban Economics*, 2017, *102*, 76–90.

Gara, Antoine, "Lennar And CalAtlantic Strike $5.7 Billion Merger To Create Nation's Largest Homebuilder," *Forbes*, October 30 2017.

Glaeser, Edward L, Joseph Gyourko, and Albert Saiz, "Housing supply and housing bubbles," *Journal of Urban Economics*, 2008, *64* (2), 198–217.

Google Finance API, "iShares Dow Jones US Home Const. (ETF)," 2018.

_ , "S&P 500 Index," 2018.

Gopal, Prashant, "As Growth in Apartment Rents Slows, U.S. Developers Press Pause," *Bloomberg News*, July 6 2016.

Graham, John R and Hyunseob Kim, "The effects of the length of the tax-loss carryback period on tax receipts and corporate marginal tax rates," Technical Report, National Bureau of Economic Research 2009.

Green, Richard K, Stephen Malpezzi, and Stephen K Mayo, "Metropolitan-specific estimates of the price elasticity of supply of housing, and their sources," *American Economic Review*, 2005, pp. 334–339.

Grenadier, Steven R, "The strategic exercise of options: Development cascades and overbuilding in real estate markets," *The Journal of Finance*, 1996, *51* (5), 1653–1679.

_ , "Option exercise games: An application to the equilibrium investment strategies of firms," *Review of financial studies*, 2002, *15* (3), 691–721.

Grimes, Arthur and Andrew Aitken, "Housing supply, land costs and price adjustment," *Real Estate Economics*, 2010, *38* (2), 325–353.

Guerrieri, V and Harald Uhlig, "Housing and credit markets: booms and busts," in "Handbook of Macroeconomics," Vol. 2, Elsevier, 2016, pp. 1427–1496.

Gutiérrez, Germán and Thomas Philippon, "Declining Competition and Investment in the US," 2017.

Gyourko, Joseph, "Housing supply," *Annu. Rev. Econ.*, 2009, *1* (1), 295–318.

_ **and Albert Saiz**, "Construction costs and the supply of housing structure," *Journal of regional Science*, 2006, *46* (4), 661–680.

He, Zhiguo and Péter Kondor, "Inefficient investment waves," *Econometrica*, 2016, *84* (2), 735–780.

Hsieh, Chang-Tai and Enrico Moretti, "Housing Constraints and Spatial Misallocation," Working Paper 21154, National Bureau of Economic Research May 2015.

Hudson, Kris, "Economists Assign Blame For Housing Shortage," *The Wall Street Journal*, June 26 2015.

Hurst, Erik and Frank P Stafford, "Home is where the equity is: Mortgage refinancing and household consumption," *Journal of money, Credit, and Banking*, 2004, *36* (6), 985–1014.

Igan, Deniz and Heedon Kang, "Do Loan-To-Value and Debt-To-Income Limits Work? Evidence From Korea," IMF Working Papers 11/297, International Monetary Fund 2011.

Kalugina, Anastasia, "Affordable Housing Policies: An Overview," *Cornell Real Estate Review*, 2016, *14* (1), 10.

Kessler, Aaron, "WCI Communities files for bankruptcy," *Sarasota Herald-Tribune*, August 5 2008.

Khouri, Andrew, "With merger closed, home builders Standard Pacific and Ryland to cut jobs," *Los Angeles Times*, October 1 2015.

Kingsley, G Thomas and Barika Xaviera Williams, *Policies for Affordable Housing in the District of Columbia: Lessons from Other Cities*, Urban Institute, 2007.

Lahart, Justin, "Fewer Home Builders Means Happier Home Builders," *Wall Street Journal*, May 15 2017.

Landvoigt, Tim, Monika Piazzesi, and Martin Schneider, "The housing market (s) of San Diego," *American Economic Review*, 2015, *105* (4), 1371–1407.

Lane, Ben, "D.R. Horton completes acquisition of real estate developer Forestar Group," *Housing Wire*, October 6 2017.

_ , "Lennar finalizes $643 million acquisition of WCI Communities," *Housing Wire*, February 10 2017.

Leamer, Edward E, "Housing is the business cycle," Technical Report, National Bureau of Economic Research 2007.

_ , "Housing really is the business cycle: what survives the lessons of 2008–09?," *Journal of Money, Credit and Banking*, 2015, *47* (S1), 43–50.

Lim, Cheng Hoon, Alejo Costa, Francesco Columba, Piyabha Kongsamut, Akira Otani, Mustafa Saiyid, Torsten Wezel, and Xiaoyong Wu, "Macroprudential policy: what instruments and how to use them? Lessons from country experiences," *IMF working papers*, 2011, pp. 1–85.

Liu, Crocker H, Adam Nowak, and Stuart S Rosenthal, "Housing price bubbles, new supply, and within-city dynamics," *Journal of Urban Economics*, 2016, *96*, 55–72.

Loecker, Jan De and Jan Eeckhout, "The rise of market power and the macroeconomic implications," Technical Report, National Bureau of Economic Research 2017.

MacDonald, Graham, *The Effect of Local Government Policies on Housing Supply*, Terner Center for Housing Innovation at UC Berkeley, 2016.

Martín, Carlos and Stephen Whitlow, "The State of the Residential Construction Industry," Technical Report, Bipartisan Policy Center September 2012.

Mayer, Christopher J and C Tsuriel Somerville, "Land use regulation and new construction," *Regional Science and Urban Economics*, 2000, *30* (6), 639–662.

_ and _ , "Residential construction: Using the urban growth model to estimate housing supply," *Journal of urban economics*, 2000, *48* (1), 85–109.

McGraw Hill Construction, "Staying Competitive in Today's Homebuilding Industry," Technical Report, Deloitte 2006.

McKenzie, Brian, "County-to-County Commuting Flows: 2006-10," Technical Report, U.S. Census Bureau 2013.

McKinsey Global Institute, "Reinventing construction: a route to higher productivity," Technical Report, McKinsey & Company February 2017.

Merle, Renae, "Caruso Homes Enters Bankruptcy Proceedings," *Washington Post*, June 25 2008.

Metcalf, Gabriel, "Sand Castles Before the Tide? Affordable Housing in Expensive Cities," *Journal of Economic Perspectives*, 2018, *32* (1), 59–80.

Metrostudy, "Metrostudy Proprietary Housing Data," 2018.

Mian, Atif and Amir Sufi, "Who bears the cost of recessions? The role of house prices and household debt," in "Handbook of Macroeconomics," Vol. 2, Elsevier, 2016, pp. 255–296.

_ , Kamalesh Rao, and Amir Sufi, "Household balance sheets, consumption, and the economic slump," *The Quarterly Journal of Economics*, 2013, *128* (4), 1687–1726.

Mirabella, Lorraine, "Home builder Gemcraft files for bankruptcy," *Baltimore Sun*, November 13 2009.

Mueller, Glenn R, "Understanding real estate's physical and financial market cycles," *Real Estate Finance*, 1995, *12*, 47–52.

Murphy, Alvin, "A Dynamic Model of Housing Supply," July 13 2015. Working paper distributed online through SSRN Electronic Journal.

Nechyba, Thomas J and Randall P Walsh, "Urban sprawl," *Journal of economic perspectives*, 2004, *18* (4), 177–200.

O'Connell, Jonathan, "Inside the rush to build Washington apartments, early signs of a bubble," *Washington Post*, June 12 2011.

O'Hollaren, Kelsey, "Home Builders in the US," Technical Report, IBISWorld February 2017.

Paciorek, Andrew, "Supply constraints and housing market dynamics," *Journal of Urban Economics*, 2013, *77*, 11–26.

Piazzesi, Monika and Martin Schneider, "Housing and macroeconomics," *Handbook of Macroeconomics*, 2016, *2*, 1547–1640.

_ , _ , and Johannes Stroebel, "Segmented housing search," Technical Report, National Bureau of Economic Research 2015.

Porter, M, "The US homebuilding industry and the competitive position of large builders," in "Centex Investor Conference, New York" 2003.

Poterba, James M, "Tax subsidies to owner-occupied housing: an asset-market approach," *The quarterly journal of economics*, 1984, pp. 729–752.

Rozenfeld, Hernán D, Diego Rybski, Xavier Gabaix, and Hernán A Makse, "The area and population of cities: New insights from a different perspective on cities," *The American Economic Review*, 2011, *101* (5), 2205–2225.

RSMeans, "RSMeans Online Construction Cost Estimating Data," 2016.

Saiz, Albert, "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Sandler, Lisa, "Developers Are Rushing to Build In Long-Shunned Los Angeles," *The Wall Street Journal*, July 26 2000.

Schnurman, Mitchell, "Why Victory Park Was a Bust," *D Magazine*, January – February 2010.

Sorkin, Andrew Ross, "Home Builders to Merge in $2.7 Billion Deal," *New York Times Dealbook*, November 4 2013.

Spatt, Chester S and Frederic P Sterbenz, "Warrant exercise, dividends, and reinvestment policy," *The Journal of Finance*, 1988, *43* (2), 493–506.

Stock, James H. and Motohiro Yogo, "Testing for weak instruments in linear IV regression," in Donald W. K. Andrews and James H. Stock, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge University Press, 2005, chapter 5.

Thompson, Boyce, "Builders Gone Bust," *Builder*, January 19 2009.

Topel, Robert and Sherwin Rosen, "Housing investment in the United States," *The Journal of Political Economy*, 1988, pp. 718–740.

U.S. Census Bureau, "American Community Survey," 2015.

_ , "Zipcode Business Patterns," 2015.

_ , "New Residential Construction," 2017.

_ , "Value of Construction Put in Place Survey," 2017.

_ , "Housing Inventory Estimate: Vacant Housing Units for Sale for the United States," 2018.

U.S. Department of Justice and Federal Trade Commission, "Horizontal Merger Guidelines," Technical Report, U.S. Department of Justice August 19 2010.

Zillow Research, "Median Home Value Per Sq Ft," 2018.

A Proofs of theoretical results

A.1 Proof of Lemma 1

To show that the solution is positive, it is sufficient to show that $(K+5n+3)\left(2Kn^2-Kn-K-3n^2-2n+1\right)$ is negative. Then, the term under the radical is less than the term outside the radical. Note that this term factors to $(K+5n+3)(n-1)(2Kn+K-3n-1)$. Under Assumption Under Assumption 1 and given that $n \geq 1$, this expression is strictly positive.

A.2 Proof of Lemma 2

From Equation 5, it is clear that m_{-}^{*} is always bounded above by $\frac{5n^2+(5K-3)n+2(K-1)}{2(K+5n+3)}$. Dividing by n and taking the limit for arbitrarily large n yields the desired result.

A.3 Proof of Proposition 1

Differentiate the term on the second line of Equation 4 that is used to define m_{-}^{*} in Equation 5 and rearrange:

$$\frac{\partial m_{-}^{*}}{\partial n}=1+\frac{7Kn-3n+K+7Km_{-}^{*}+9m_{-}^{*}}{5n^2-5Kn+3n+2K+2-10m_{-}^{*}n-2m_{-}^{*}K-6m_{-}^{*}} \tag{12}$$

From the chain rule, the sign of $\frac{\partial\left(m_{-}^{*}/n\right)}{\partial n}$ is the same as the sign of $\frac{\partial m_{-}^{*}}{\partial n}-\frac{m_{-}^{*}}{n}$. Accordingly, to show that the latter is positive, it remains only to show that the fraction in Equation 12 is positive. Under Assumption 1, the numerator is positive. Rearranging the denominator yields the following condition on m_{-}^{*} for the denominator to be positive:

$$m_{-}^{*}<\frac{5n^2+5Kn+3n-2K+2}{2(K+5n+3)} \tag{13}$$

This is exactly the condition implied by Lemma 1 for m^*_-. Under Assumption 1, m^* as specified by Equation 5 satisfies this restriction. This completes the proof.

A.4 Proof of Proposition 2

From Lemma 1 and 2, $\frac{m^*_-}{n}$ ranges from zero to $\frac{1}{2}$ in the large n limit. From Proposition 1, $\frac{m^*_-}{n}$ is continuously increasing in n. Therefore, by the intermediate value theorem, at any n the value of $\frac{m^*_-}{n}$ lies between zero and $\frac{1}{2}$ — i.e., $m^* \in (0, n-1)$.

A.5 Proof of Proposition 3

Differentiate H_1 as specified in Equation 7 using $m' = \frac{\partial m}{\partial n}$ for notational clarity:

$$\frac{\partial H_1}{\partial n} = \frac{K+1}{\alpha_1 + \beta_1} \frac{2n^2 m' - 2nm' - m^2 - 2m}{\left(m+2\right)^2 \left(n-1\right)^2} \tag{14}$$

It remains to show that $2n^2m' - 2nm' - m^2 - 2m$ is positive. To see this, note that Lemma 1 shows that $m' > 1$ and Lemma 2 and Proposition 2 show that $m \leq \frac{n}{2}$. From this, it follows that $2n^2m' - 2nm' - m^2 - 2m \geq 2n\left(n-1\right)$ for $n > 1$. Since this term is positive, the expression on the right-hand side of Equation 14 is also positive.

A.6 Proof of Proposition 4

Let $\mathrm{SD}\left(p_1\right)$ denote the *ex post* standard deviation in the realization of p_1. Since n firms are each building the quantity h_1 specified by Equation 6 with probability $\frac{m}{n-1}$ and given the price at $t = 1$ as specified by Equation 1, $\mathrm{SD}\left(p_1\right)$ may be written in terms of m as follows:

$$\mathrm{SD}\left(p_1\right) = \frac{\alpha_1\left(K+1\right)}{\alpha_1 + \beta_1} \frac{nm\left(n-m-1\right)}{\left(n-1\right)\left(m+2\right)} \tag{15}$$

Differentiating Equation 15 with respect to n and rearranging yields the following result:

$$\mathrm{sign}\left(\frac{\partial}{\partial n}\mathrm{SD}\left(p_1\right)\right) = \mathrm{sign}\left(n\left(n-1\right)m'\left(2n - mn - 3m - 2\right) + m\left(m+2\right)\left(mn - n + m + 1\right)\right) \tag{16}$$

The term $\left(2n - mn - 3m - 2\right)$ is positive when $m < \frac{2n(n-1)}{n^2+3}$ and the term $\left(mn - n + m + 1\right)$ is positive when $m > \frac{n(n-1)}{n(n+1)}$.

FOR ONLINE PUBLICATION

B Tests of robustness

B.1 Results without low-population markets

Table A1: Regression results where the dependent variable is the logarithm of the total value of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.029	1.20	0.016	1.62	0.029	1.20
	(0.015)	(0.18)	(0.016)	(0.28)	(0.015)	(0.18)
Jobs within 50 miles	-5.44	6.53	-5.76	7.03	-5.44	6.53
	(0.60)	(1.97)	(0.60)	(2.44)	(0.60)	(1.97)
Construction cost	-0.47	-0.56	-0.39	-0.41	-0.47	-0.56
	(0.036)	(0.050)	(0.036)	(0.056)	(0.036)	(0.050)
Share national firms			0.70	-0.099		
			(0.099)	(0.21)		
Share regional firms			-0.17	-0.22		
			(0.11)	(0.17)		
Share micro firms			0.22	-2.10		
			(0.088)	(0.43)		
Established market					1.08	
					(0.25)	
Observations	11521	11521	11521	11521	11521	11521
R^2	0.773	0.579	0.777	0.465	0.773	0.579
1^{st} Stage F-statistic		95.948		54.514		95.948
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A2: Regression results where the dependent variable is the logarithm of the total square footage of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.043	1.19	0.026	1.60	0.043	1.19
	(0.014)	(0.17)	(0.015)	(0.27)	(0.014)	(0.17)
Jobs within 50 miles	-4.26	7.48	-4.73	7.83	-4.26	7.48
	(0.56)	(1.88)	(0.55)	(2.34)	(0.56)	(1.88)
Construction cost	-0.46	-0.54	-0.37	-0.39	-0.46	-0.54
	(0.033)	(0.048)	(0.034)	(0.054)	(0.033)	(0.048)
Share national firms			0.58	-0.20		
			(0.092)	(0.20)		
Share regional firms			-0.55	-0.59		
			(0.11)	(0.17)		
Share micro firms			0.067	-2.21		
			(0.082)	(0.41)		
Established market					-3.76	
					(0.59)	
Observations	11516	11516	11516	11516	11516	11516
R^2	0.763	0.538	0.770	0.408	0.763	0.538
1^{st} Stage F-statistic		95.906		54.470		95.906
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A3: Regression results where the dependent variable is the logarithm of the absolute value of the change in the number of the units of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.39	2.01	0.32	2.40	0.39	2.01
	(0.042)	(0.39)	(0.045)	(0.58)	(0.042)	(0.39)
Jobs within 50 miles	5.14	21.7	4.61	21.2	5.14	21.7
	(1.67)	(4.35)	(1.66)	(4.96)	(1.67)	(4.35)
Construction cost	-1.22	-1.34	-0.99	-1.01	-1.22	-1.34
	(0.099)	(0.11)	(0.10)	(0.11)	(0.099)	(0.11)
Share national firms			2.57	1.54		
			(0.28)	(0.43)		
Share regional firms			0.15	0.094		
			(0.31)	(0.35)		
Share micro firms			1.32	-1.68		
			(0.25)	(0.88)		
Established market					-3.87	
					(0.70)	
Observations	11493	11493	11493	11493	11493	11493
R^2	0.339	0.198	0.353	0.154	0.339	0.198
1^{st} Stage F-statistic		99.243		55.233		99.243
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A4: Regression results where the dependent variable is the months of supply of housing (including under construction, model and finished vacant) at contemporaneous absorption rates.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.24	1.49	0.21	1.90	0.24	1.49
	(0.063)	(0.28)	(0.067)	(0.40)	(0.063)	(0.28)
Jobs within 50 miles	11.9	9.60	11.0	8.48	11.9	9.60
	(5.51)	(6.13)	(5.53)	(6.60)	(5.51)	(6.13)
Construction cost	-6.05	-1.25	-5.91	0.79	-6.05	-1.25
	(0.78)	(1.35)	(0.79)	(1.81)	(0.78)	(1.35)
Share national firms			0.67	0.44		
			(0.41)	(0.49)		
Share regional firms			-0.071	-0.38		
			(0.41)	(0.49)		
Share micro firms			0.62	-1.10		
			(0.34)	(0.57)		
Established market					-13.5	
					(3.25)	
Observations	3248	3248	3248	3248	3248	3248
R^2	0.398	0.231	0.402	0.121	0.398	0.231
1st Stage F-statistic		95.703		56.954		95.703
1st Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A5: Regression results where the dependent variable is the logarithm of the absolute value of the change in price per square foot.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	-0.16	-4.52	-0.098	-6.10	-0.16	-4.52
	(0.045)	(0.59)	(0.049)	(0.98)	(0.045)	(0.59)
Jobs within 50 miles	-12.9	-57.7	-11.6	-59.5	-12.9	-57.7
	(1.80)	(6.61)	(1.78)	(8.41)	(1.80)	(6.61)
Construction cost	-0.27	0.039	-0.57	-0.51	-0.27	0.039
	(0.11)	(0.17)	(0.11)	(0.19)	(0.11)	(0.17)
Share national firms			-2.53	0.47		
			(0.30)	(0.72)		
Share regional firms			0.97	1.17		
			(0.34)	(0.61)		
Share micro firms			-0.76	7.97		
			(0.26)	(1.50)		
Established market					-15.2	
					(1.91)	
Observations	11511	11511	11511	11511	11511	11511
R^2	0.171	-0.943	0.194	-1.599	0.171	-0.943
1^{st} Stage F-statistic		95.863		54.244		95.863
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

B.2 Results without low-production markets

Table A6: Regression results where the dependent variable is the logarithm of the total value of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.051	1.15	0.046	1.44	0.051	1.15
	(0.013)	(0.18)	(0.014)	(0.27)	(0.013)	(0.18)
Jobs within 50 miles	0.43	7.63	0.16	7.52	0.43	7.63
	(0.48)	(1.35)	(0.48)	(1.57)	(0.48)	(1.35)
Construction cost	-0.32	-0.47	-0.25	-0.37	-0.32	-0.47
	(0.033)	(0.050)	(0.033)	(0.052)	(0.033)	(0.050)
Share national firms			0.99	0.58		
			(0.083)	(0.14)		
Share regional firms			0.23	0.22		
			(0.091)	(0.13)		
Share micro firms			0.45	-1.17		
			(0.074)	(0.33)		
Established market					-0.0071	
					(1.45)	
Observations	15727	15727	15727	15727	15727	15727
R^2	0.785	0.630	0.790	0.568	0.785	0.630
1^{st} Stage F-statistic		85.312		52.567		85.312
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A7: Regression results where the dependent variable is the logarithm of the total square footage of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.048	1.10	0.042	1.36	0.048	1.10
	(0.012)	(0.18)	(0.013)	(0.25)	(0.012)	(0.18)
Jobs within 50 miles	0.55	7.41	0.28	7.24	0.55	7.41
	(0.46)	(1.28)	(0.46)	(1.48)	(0.46)	(1.28)
Construction cost	-0.33	-0.47	-0.26	-0.37	-0.33	-0.47
	(0.031)	(0.047)	(0.031)	(0.049)	(0.031)	(0.047)
Share national firms			0.98	0.57		
			(0.079)	(0.14)		
Share regional firms			0.17	0.13		
			(0.088)	(0.13)		
Share micro firms			0.44	-1.11		
			(0.071)	(0.31)		
Established market					1.13	
					(0.51)	
Observations	15714	15714	15714	15714	15714	15714
R^2	0.772	0.602	0.777	0.539	0.772	0.602
1^{st} Stage F-statistic		84.764		52.443		84.764
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A8: Regression results where the dependent variable is the logarithm of the absolute value of the change in the number of the units of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.17	1.49	0.18	1.79	0.17	1.49
	(0.034)	(0.39)	(0.036)	(0.54)	(0.034)	(0.39)
Jobs within 50 miles	-1.70	6.95	-2.69	5.90	-1.70	6.95
	(1.27)	(2.89)	(1.26)	(3.17)	(1.27)	(2.89)
Construction cost	-1.03	-1.22	-0.85	-0.99	-1.03	-1.22
	(0.086)	(0.11)	(0.086)	(0.11)	(0.086)	(0.11)
Share national firms			2.04	1.49		
			(0.22)	(0.30)		
Share regional firms			-0.15	-0.21		
			(0.24)	(0.26)		
Share micro firms			0.45	-1.47		
			(0.20)	(0.67)		
Established market					8.42	
					(3.79)	
Observations	15651	15651	15651	15651	15651	15651
R^2	0.321	0.218	0.334	0.196	0.321	0.218
1^{st} Stage F-statistic		86.219		51.970		86.219
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A9: Regression results where the dependent variable is the months of supply of housing (including under construction, model and finished vacant) at contemporaneous absorption rates.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.19	2.06	0.20	3.30	0.19	2.06
	(0.048)	(0.51)	(0.051)	(0.99)	(0.048)	(0.51)
Jobs within 50 miles	0.071	-7.00	0.30	-11.6	0.071	-7.00
	(3.81)	(5.28)	(3.81)	(7.30)	(3.81)	(5.28)
Construction cost	-7.05	-0.23	-6.82	3.26	-7.05	-0.23
	(0.61)	(2.00)	(0.62)	(3.36)	(0.61)	(2.00)
Share national firms			0.37	2.10		
			(0.29)	(0.73)		
Share regional firms			-0.36	2.12		
			(0.30)	(0.93)		
Share micro firms			-0.035	-1.28		
			(0.24)	(0.56)		
Established market					-18.6	
					(3.69)	
Observations	4374	4374	4374	4374	4374	4374
R^2	0.406	-0.031	0.410	-0.651	0.406	-0.031
1^{st} Stage F-statistic		31.336		14.738		31.336
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A10: Regression results where the dependent variable is the logarithm of the absolute value of the change in price per square foot.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	-0.12	-3.24	-0.034	-3.94	-0.12	-3.24
	(0.037)	(0.52)	(0.038)	(0.75)	(0.037)	(0.52)
Jobs within 50 miles	-8.78	-29.1	-8.61	-29.2	-8.78	-29.1
	(1.36)	(3.81)	(1.36)	(4.40)	(1.36)	(3.81)
Construction cost	-0.23	0.19	-0.34	-0.017	-0.23	0.19
	(0.092)	(0.14)	(0.093)	(0.15)	(0.092)	(0.14)
Share national firms			-2.09	-0.89		
			(0.24)	(0.41)		
Share regional firms			-0.43	-0.29		
			(0.26)	(0.37)		
Share micro firms			-1.62	2.96		
			(0.21)	(0.93)		
Established market					-11.5	
					(1.49)	
Observations	15709	15709	15709	15709	15709	15709
R^2	0.180	-0.428	0.191	-0.655	0.180	-0.428
1^{st} Stage F-statistic		84.742		52.290		84.742
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

B.3 Results with adjacent markets removed from the instrument

Table A11: Regression results where the dependent variable is the logarithm of the total value of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.042	1.21	0.029	1.38	0.042	1.21
	(0.013)	(0.21)	(0.014)	(0.27)	(0.013)	(0.21)
Jobs within 50 miles	1.50	8.88	1.53	8.26	1.50	8.88
	(0.49)	(1.48)	(0.49)	(1.49)	(0.49)	(1.48)
Construction cost	-0.32	-0.51	-0.29	-0.42	-0.32	-0.51
	(0.033)	(0.056)	(0.034)	(0.054)	(0.033)	(0.056)
Share national firms			0.66	0.39		
			(0.087)	(0.13)		
Share regional firms			0.38	0.25		
			(0.094)	(0.13)		
Share micro firms			0.51	-0.99		
			(0.077)	(0.31)		
Established market					-3.19	
					(1.46)	
Observations	14827	14827	14827	14827	14827	14827
R^2	0.793	0.621	0.794	0.588	0.793	0.621
1^{st} Stage F-statistic		67.808		51.306		67.808
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A12: Regression results where the dependent variable is the logarithm of the total square footage of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.038	1.41	0.026	1.62	0.038	1.41
	(0.013)	(0.22)	(0.013)	(0.29)	(0.013)	(0.22)
Jobs within 50 miles	1.42	9.95	1.40	9.30	1.42	9.95
	(0.47)	(1.55)	(0.47)	(1.59)	(0.47)	(1.55)
Construction cost	-0.33	-0.55	-0.31	-0.46	-0.33	-0.55
	(0.032)	(0.059)	(0.032)	(0.058)	(0.032)	(0.059)
Share national firms			0.56	0.22		
			(0.083)	(0.14)		
Share regional firms			0.21	0.018		
			(0.091)	(0.15)		
Share micro firms			0.41	-1.38		
			(0.074)	(0.34)		
Established market					-2.97	
					(1.40)	
Observations	14817	14817	14817	14817	14817	14817
R^2	0.776	0.498	0.777	0.437	0.776	0.498
1^{st} Stage F-statistic		67.517		51.194		67.517
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A13: Regression results where the dependent variable is the logarithm of the absolute value of the change in the number of the units of housing produced.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.27	0.33	0.31	0.20	0.27	0.33
	(0.036)	(0.42)	(0.038)	(0.50)	(0.036)	(0.42)
Jobs within 50 miles	-1.71	-1.33	-2.64	-3.19	-1.71	-1.33
	(1.32)	(2.98)	(1.32)	(2.84)	(1.32)	(2.98)
Construction cost	-1.09	-1.10	-0.90	-0.89	-1.09	-1.10
	(0.090)	(0.11)	(0.091)	(0.10)	(0.090)	(0.11)
Share national firms			2.16	2.18		
			(0.23)	(0.25)		
Share regional firms			0.14	0.15		
			(0.25)	(0.25)		
Share micro firms			0.45	0.57		
			(0.21)	(0.60)		
Established market					8.76	
					(3.96)	
Observations	14783	14783	14783	14783	14783	14783
R^2	0.324	0.324	0.336	0.335	0.324	0.324
1^{st} Stage F-statistic		67.128		51.530		67.128
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A14: Regression results where the dependent variable is the months of supply of housing (including under construction, model and finished vacant) at contemporaneous absorption rates.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	0.23	0.97	0.25	1.20	0.23	0.97
	(0.044)	(0.50)	(0.046)	(0.61)	(0.044)	(0.50)
Jobs within 50 miles	-12.0	-17.6	-11.5	-18.2	-12.0	-17.6
	(3.50)	(5.27)	(3.51)	(5.69)	(3.50)	(5.27)
Construction cost	-7.87	-5.08	-7.68	-4.47	-7.87	-5.08
	(0.55)	(1.96)	(0.55)	(2.13)	(0.55)	(1.96)
Share national firms			0.17	0.61		
			(0.27)	(0.40)		
Share regional firms			-0.33	0.30		
			(0.26)	(0.49)		
Share micro firms			-0.25	-0.70		
			(0.22)	(0.37)		
Established market					64.7	
					(9.60)	
Observations	4030	4030	4030	4030	4030	4030
R^2	0.502	0.427	0.505	0.397	0.502	0.427
1^{st} Stage F-statistic		16.148		13.203		16.148
1^{st} Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.

Table A15: Regression results where the dependent variable is the logarithm of the absolute value of the change in price per square foot.

	OLS	IV	OLS	IV	OLS	IV
Firms producing 90%	-0.055	-5.26	0.011	-6.02	-0.055	-5.26
	(0.038)	(0.77)	(0.040)	(1.00)	(0.038)	(0.77)
Jobs within 50 miles	-11.8	-44.3	-11.7	-41.6	-11.8	-44.3
	(1.40)	(5.38)	(1.39)	(5.55)	(1.40)	(5.38)
Construction cost	-0.30	0.53	-0.42	0.15	-0.30	0.53
	(0.095)	(0.21)	(0.096)	(0.20)	(0.095)	(0.21)
Share national firms			-2.35	-1.06		
			(0.25)	(0.50)		
Share regional firms			-0.59	0.15		
			(0.27)	(0.52)		
Share micro firms			-1.74	5.03		
			(0.22)	(1.19)		
Established market					35.7	
					(4.18)	
Observations	14812	14812	14812	14812	14812	14812
R^2	0.187	-1.448	0.199	-1.778	0.187	-1.448
1st Stage F-statistic		67.494		51.172		67.494
1st Stage p-value		0.000		0.000		0.000

Standard errors in parentheses.

All specifications include firm, year, and market fixed effects.