Online Appendix - Boom Town Business Dynamics

Ryan A. Decker, Meagan McCollum, Gregory B. Upton, Jr.

Journal of Human Resources

A Model

In this section we briefly sketch illustrative model intuition behind our theoretical framework. Clementi and Palazzo (2016) construct a fully featured model of firm dynamics for studying the cyclical properties of business entry. Here we describe a simplified version of that model to explore key results. The differences between our model here and that of Clementi and Palazzo (2016) are (a) we omit capital from the model and (b) we study a simple transition path exercise rather than implementing full stochastic aggregate risk and business cycle exercises. We also initially differ by shutting down ex ante heterogeneity of entrants, but we expand our investigation to include ex ante heterogeneity further below. While we do calibrate the model, we take much of our calibration from existing literature and focus primarily on the qualitative results.

Firms face idiosyncratic productivity draws z and an aggregate productivity state A. Idiosyncratic productivity evolves according to $\ln z' = \rho_z \ln z + \sigma_z \varepsilon_z'$ where $\varepsilon_z \sim N(0,1)$; this yields a conditional distribution of z' given by $H(z'|z)$. Firms produce using technology Azn^{α}, where α governs revenue curvature (which we interpret here as decreasing returns to scale); firms discount profits with factor β and face a spot market for labor with wage w and labor supply curve $L_s(w) = w^{\gamma}$ (with $\gamma > 0$). Continuing firms must pay a fixed operating cost c_f; the operating cost is not persistent, and $c_f \sim LN(\mu_c, \sigma_c)$.

Under these assumptions, entry is determined by the free entry condition:

$$c_e = \mathbb{E}_{z'} V(z'; A, w), \tag{9}$$

where c_e is the entry cost and $V(z'; A, w)$ is the value function of an operating firm. The mass of entrants is determined in equilibrium such that the average firm value is pinned down to the entry cost. Upon entry, new entrants receive productivity draws consistent with the unconditional productivity distribution of incumbent firms.

The timing of the model is as follows. At the beginning of the period, incumbents observe their productivity z then hire labor and produce. Incumbents, following production, draw their operating cost c_f then choose whether to continue or exit; at the same time, the mass of entrants is determined, and entrants pay the entry cost c_e. Then the next period begins.

The incumbents' problem is as follows. First, the incumbent faces a static profit maximization problem yielding the following first-order condition for labor demand:

$$n(z; A, w) = \left(\frac{w}{\alpha A z}\right)^{\frac{1}{\alpha-1}}. \tag{10}$$

This yields a profit function $\pi(z; A, w)$.

Here we can illustrate the central intuition of incumbent behavior in the model. Suppose A increases by x percent. Our interest is in the growth of firm-level employment in a comparative statics view:

$$g = \frac{n(z; (1+x)A, w) - n(z; A, w)}{n(z; A, w)} \tag{11}$$

where we hold the wage constant (i.e., partial equilibrium). In this simple environment it is straightforward to show that

$$g = (1+x)^{\frac{1}{1-\alpha}} - 1, \tag{12}$$

that is, the firm's employment growth response is a function only of x and α. Importantly for our study, the absolute value of the growth rate is increasing in α or, equivalently, decreasing in the curvature of the revenue function. Revenue function curvature dampens the response of incumbents to shocks.[33] Equivalently, revenue function curvature compresses the distribution of labor demand across firms of different productivity realizations. This creates opportunity for aggregate shocks to affect the number of firms, not just the size of preexisting firms.

Returning to the model environment, the value of an incumbent at the beginning of a period is given by:

$$V(z; A, w) = \pi(z; A, w) + \beta \mathbb{E}_{c_f}\left[\max\{0, \mathbb{E}_{z'|z} V(z'; A, w) - c_f\}\right] \tag{13}$$

This optimization problem yields an exit rule such that firms choose to exit when the expected value of the firm is negative (where exit provides a payoff of zero, as shown in the internal maximization operator); this results in a threshold rule such that incumbents exit when $z \leq z^*(A, w)$.

The recursive competitive equilibrium is defined as follows. $V(z; A, w), n(z; A, w)$, and the

[33] Decker et al. (Forthcoming) show that this result holds even in a more fully specified and calibrated model with labor adjustment costs.

associated exit rule arising from the threshold z^* solve the incumbents' problem, and the mass of entrants M is such that the free entry condition (9) holds with equality; the distribution of new entrants is given by $E(z') = M^* H(z')$. The labor market clears; that is, $w^\gamma = \int n(z; A, w) d\Gamma(z)$, where $\Gamma(z)$ is the measure of producing firms (distributed over z). Finally, the measure of firms evolves according to $\Gamma'(z') = \int \int_{c_f} \int_{z^*}^{\infty} d\Gamma(z) dG(c_f) dH(z'|z) + E(z')$. The latter condition simply illustrates that the new distribution of firms reflects the distribution of incumbents that chose not to exit, appropriately transitioned to updated productivity draws, plus the mass and distribution of new entrants.

We calibrate the model as reported on Table A1 in the column labeled "Model 1"; this calibration mostly follows Clementi and Palazzo (2016) except that we choose μ_c (the operating cost distribution mean) to target an entry rate of 9% (that is, entrants account for 9% of firms), consistent with Business Dynamics Statistics data from the early 2000s.

We solve the steady state of the model by starting with guesses for the entry mass M and the wage w, solving value functions and policy functions (via value function iteration), iterating to a stationary distribution where $\Gamma' = \Gamma$, checking labor market clearing, revising the wage until the market clears, then revising the entry mass M until the free entry condition holds. We consider two steady states; in the baseline steady state we set $A = 1$, and in the expansion steady state we set $A = 1.2$ (these choices are arbitrary, designed only to illustrate qualitative dynamics). We then study a transition from the baseline to the expansion state. In period 0, the economy is in the baseline steady state with no expectation for change. In period 1, firms learn that A will transition from 1 to 1.2 effective the beginning of period 2, after which the economy will converge to the steady state associated with $A = 1.2$ and no expectation of change. This exercise is illustrated on the top left panel of Figure A1. The positive aggregate shock (solid line) causes a permanent increase in the number of firms (short-dashed line); this rise in the firm count is facilitated by a surge in entry (long-dashed line), including the employment-weighted entry rate (dot-dashed line). This result (surging entry and employment-weighted entry) is robust to a wide range of parameterizations.

We next generalize the model slightly to allow ex ante heterogeneity among entrants. At any time there exists a mass M_p of potential entrants. Each potential entrant receives a *signal* about their productivity given by $q \sim Pareto(\min(z), \xi)$. The signal q relates to productivity on entry with the conditional distribution $H(z'|q)$; that is, productivity on entry follows $\ln z' = \rho_z \ln q + \sigma_z \varepsilon'_z$. While it is not strictly necessary that the distribution of potential entrants' signals differ from the distribution of incumbents' productivity, doing so makes it possible to match the number and size of entrant firms to the data. While incumbent firms are producing, potential entrants observe their signal q and choose whether to enter for production in the next period. The potential entrants' problem is solved simply by choosing to enter when the free entry condition holds:

$$\beta \mathbb{E}_{z'|q} V(z'; A, w) \geq c_e \tag{14}$$

As is common in models of this class, this free entry condition yields an entry rule such that potential entrants choose to enter if and only if $q \geq q^*(A, w)$, where $q^*(A, w)$ is a threshold value dependent on the aggregate state. This threshold rule differs in important ways from the simpler free entry condition given by (9); in particular, the threshold rule does not hold with equality and, therefore, has less stark implications for the value of existing firms. Additionally, the productivity distribution of new entrants differs from that of continuing incumbents due to the signal distribution; this is necessary for matching the firm size distribution (as noted by Clementi and Palazzo (2016)), but it creates different dynamics for the employment share of entrants. On Table A1, the column "Model 2" reports calibration details for this model generalization.

We conduct the same transition path exercise as above, reported on the top right panel of Figure A1. The solid line reports the path of aggregate productivity. The short-dashed line shows that, as in the previous experiment, the improvement in aggregate conditions causes a rise in the number of firms as existing firms do not grow enough to accommodate the shock. The long-dashed line shows that, in this calibration, the rise in the number of firms is facilitated in part by a surge in entry. However, unlike the previous experiment, the dot-dashed line shows that the employment share of entrants does not rise. This is the result of ex ante heterogeneity and quality signals; in this setup, the rise in entry is driven by a decline in the threshold for the productivity signal above which entry is profitable. This induces a selection mechanism in which the positive aggregate shock allows lower-quality entrepreneurs to enter; upon entering, their employment is lower than the minimum productivity of entrants during the initial stationary state.

The exercises from our more general model still support the notion that aggregate shocks are accommodated, at least in part, by a rise in entry. However, even this result is heavily influenced by calibration. For example, the bottom left panel of Figure A1 reports the same experiment except that the revenue curvature parameter α is set at 0.7 (rather than 0.8); in this experiment, even the unweighted entry rate responds negatively to the shock (note that the number of firms still rises, facilitated by a lower exit rate). The bottom right panel of Figure A1 shows that the entry rate effect can be reduced by lowering the labor supply elasticity to $\gamma = 1$ (from $\gamma = 2$). Future research might further explore calibration considerations in relation to our empirical results.

B Data

B.1 County Business Patterns

County Business Patterns (CBP) is based on the Census Bureau's Business Register and covers almost all private employer establishments in the U.S. Non-employers—those businesses without employees for Social Security Administration purposes—are excluded; however, the employer universe covers potentially all legal forms of organization, including sole proprietors (among whom are many employers). See DeSalvo et al. (2016) for details on the Business Register data underlying CBP.

CBP covers the universe of private business establishments, excluding only the following NAICS industries: 111 and 112 (crop and animal production), 482 (rail transportation), 491 (Postal Service), 525110, 525120, 525190 (pension, health, welfare, and vacation funds), 525920 (trusts, estates, and agency accounts), 814 (private households), and 92 (public administration). Government-owned businesses in the following NAICS industries are included: 4248 (wholesale liquor establishments), 44531 (retail liquor stores), 511130 (book publishers), 522120 (federally-chartered savings institutions), 522130 (federally-chartered credit unions), and 622 (hospitals).[34]

While establishment counts are published for all industry-by-county cells in CBP data, employment counts are suppressed in some cells. In these cases, a size range is reported instead of a precise employment count. We use these size range reports to impute employment to suppressed cells; we first impute employment for any suppressed county-level observations, then we impute employment for suppressed county-by-sector observations.

We impute suppressed county-level employment as follows. Within a given year, we first categorize all non-suppressed counties into size bins that correspond to the size bins reported for suppressed counties. We then obtain average actual county employment by size bin (among non-suppressed counties) and populate the employment variable for suppressed counties with the average employment of non-suppressed counties that have reported employment within the corresponding size bin.[35] That is, we estimate that suppressed counties

[34] See https://www.census.gov/programs-surveys/cbp/technical-documentation/methodology.html.

[35] If there are no counties with employment in the indicated size bin, we assign the midpoint of the size bin for all bins except the top bin, for which we assign the lower bound of the bin.

have employment equal to the average employment of non-suppressed counties in the same size class. Next, we sum up total U.S. employment for the year by adding total employment among non-suppressed counties with total estimated employment among suppressed counties. We compare this estimated total to actual reported total U.S. employment for the year (which is available in separate national-level CBP files). Observing the discrepancy between true national employment and our national estimate based on our initial imputation for suppressed counties, we then modify our estimated employment for suppressed counties by sharing out the discrepancy proportionally (based on each county's estimated share of total suppressed employment). The result is our final estimate of employment in each suppressed county. Our imputation method therefore assumes that true county employment for suppressed counties is distributed among employment size bins in a manner similar to the employment distribution of non-suppressed counties, but adjusted to ensure that county employment adds up to true national employment. Observations in which county-level employment is suppressed comprise no more than 0.2% of employment, depending on the year; prior to 2011 imputed observations never account for more than 0.1% of employment.

With populated county-level employment values in hand (whether true or imputed), we next impute employment for suppressed county-by-sector cells (where sectors are defined by two-digit NAICS codes). We proceed in a fashion similar to our county employment imputation method, but our imputation now uses sector-specific averages. Specifically: within a given year *and sector,* we obtain average employment by employment size bin among non-suppressed county-by-sector cells then apply that average to each suppressed cell according to its reported employment bin. After doing this for each sector, we add up sector employment (that is, the sum of total employment among non-suppressed cells and estimated employment among suppressed cells) *by county* and compare the estimated county employment to true reported county employment (or, in the case of counties in which county employment was suppressed, we compare to estimated total county employment as constructed above). Observing the discrepancy between total reported county employment and total county employment based on estimated sector cells, we then adjust our estimates for suppressed cells by sharing out the county-level discrepancy in manner proportional to the initial estimates. This method ensures that sector-level employment within counties adds up to total county employment appropriately. Observations in which county-by-sector employment is suppressed comprise between 1.1 and 3.2% of employment (after imputation), depending on the year (the share generally increases over time).

Finally, since some of our exercises involve narrower industry groups requiring 3-digit

NAICS aggregations, we also impute data for certain county-by-3-digit-NAICS cells (211, 213, 324, and 325). For these we proceed in similar fashion to our approach described above: by year, we estimate cell employment based on the nationwide average of cell employment for each specific 3-digit industry. We then adjust these estimates by aggregating to the county-by-sector level. That is, we adjust estimates for cells of NAICS 211 and 213 by aggregating to county NAICS 21 (mining) employment, and we adjust estimates for cells of NAICS 324 and 325 by aggregating to county NAICS 31-33 (manufacturing) employment. Note that this method requires us to determine suppression and impute for all 3-digit naics industries within these two sectors (21 and 31-33). Observations in which county-by-industry employment for the 3-digit industries in NAICS 21 and 22 is suppressed comprise between 2.6 and 3.5% of employment (after imputation), depending on the year (the share generally increases over time).

B.2 Longitudinal Business Database

The Longitudinal Business Database (LBD), like CBP, is based on the Census Bureau's Business Register. The two datasets also share the same industry scope. Jarmin and Miranda (2002) describe the construction of the LBD. Critically for our purposes, the LBD consists of establishment-level data with longitudinal establishment identifiers. The data also include a firm identifier linking establishments under common ownership or operational control; importantly, this firm identifier is superior to simple tax identifiers (i.e., EINs), since some firms have multiple EINS. Industry codes correspond to establishments. For our purposes it is not necessary to assign an industry code to firms; all industry categories are based on establishment industry (and, as such, industry characteristics of "new firms" actually reflect the industry characteristics of establishments of new firms in a given county).

Importantly, while the LBD consists of establishment-level microdata, county aggregates are the units of observation for our study. All regressions and summary statistics are based on county observations. For example, a key dependent variable is the new firm growth rate component, which is the share of county-level employment growth accounted for by the job creation of new firms in a given year.

Consistent with much of the literature (e.g., Haltiwanger et al., 2013), we define an *establishment* birth as the first year in which an establishment has positive employment, and we determine *firm* age as follows: when a firm identifier first appears in the data, it is assigned the age of its oldest establishment; thereafter, the firm ages naturally each year.[36]

[36]The establishment-level longitudinal linkages in the LBD are generally considered to be of high integrity.

When studying annual growth rates, we focus firm-level calculations on "organic growth" as in Haltiwanger et al. (2013) and subsequent literature, in which the lagged employment term emp^j_{ct-1} is comprised of the lagged employment of all establishments in county c that belong to firms in group j in year t. This approach allows us to abstract from growth driven by merger and acquisition activity. In practice this means that the growth of an establishment that changes firm owners between years $t-1$ and t is assigned to the firm that owns the establishment as of time t. In cumulative employment growth exercises, on the other hand, we make no attempt to ensure that growth is "organic" since it is not clear how to interpret organic growth in this context. As such, however, the employment share of post-2006 new firms in any give year can, in principal, include employment of establishments that are older but were acquired by those new firms during the post-2007 period.

C Additional results and robustness checks

C.1 Manufacturing

Figure 3 (discussed in Section 4.2) shows the 2006-2014 change in employment and establishment counts in "boom towns" by sector (from CBP data), where the manufacturing sector shows modestly negative growth.[37] Figure A4 reports the same variables in the control group superset, where a much larger decline in manufacturing—associated with the Great Recession—is evident. In other words, manufacturing activity in the "boom towns" fell by less than in the control group counties. The result is that our diff-in-diff exercises find positive causal effects of the shale boom on manufacturing employment growth rates, as shown on Table A5.

Since Cosgrove et al. (2015) find negative effects of the shale boom on manufacturing employment, it is worth exploring the sector more. Figure A5 reports the 2006-2014 change in employment and establishment counts in the manufacturing sector by play (and for the non-shale control group superset). While there is wide variation across plays, most shale plays see declines in manufacturing activity; Appalachia and Haynesville see larger declines than the non-shale control group counties, consistent with Cosgrove et al. (2015) who study

Unfortunately, the longitudinal linkages of the LBD's firm identifiers are less reliable and are therefore a source of measurement error. Nevertheless, we follow much recent literature in proceeding with firm age concepts that rely on the LBD firm identifier; these concepts are made more robust by the popular method, which we adopt, of assigning firm age based on establishment age at the firm's first appearance.

[37] Interestingly, Figure A8 shows that roughly half of the 2006-2014 decline in manufacturing employment can be accounted for by net establishment exit.

the Appalachia play alone. Only Bakken sees positive gains in manufacturing activity.

C.2 Agriculture

Given the rural nature of many shale counties, a natural question is how agricultural activity responded to the shale boom. Both CBP and the LBD have only partial coverage of the agricultural sector (NAICS 11); the most important limitation is that farms (i.e., NAICS 111 and 112, encompassing crop and animal production) are omitted from both data sources. Remaining industries included in CBP and the LBD are in NAICS 113 (forestry and logging), 114 (fishing, hunting, and trapping), and 115 (support activities for agriculture and forestry). With that in mind, Figure A6 shows CBP employment and establishment counts for covered agricultural industries (i.e., NAICS 11 excluding 111 and 112), scaled by average (county-level) employment and establishment counts for 2006, for shale counties overall, "boom towns," and counties in the non-shale control set. No clear pattern is evident; during the boom period, shale counties overall see a more substantial decline in employment than do the other counties, but "boom towns" see somewhat higher employment than the control areas. Establishment counts move similarly in all county groups during the boom after being somewhat higher in shale counties during the pre-boom period.

Since farms (NAICS 111 and 112) are missing from CBP data, we obtain county-level farm data from the U.S. Department of Agriculture's Census of Agriculture. These county-level data are only available at five-year intervals consistent with the census timing, so we cannot observe the exact time frame and frequency we study in other exercises, but we can observe data for 2002, 2007, 2012, and 2017, roughly corresponding with the 2000-2014 period. Farm count data are available for all counties outside Alaska, while data on hired workers (roughly equivalent to the employment concept used in CBP and the LBD) are available for almost all counties.[38]

Figure A7 shows hired labor and establishment counts relative to 2007 county averages. Patterns of hired labor are similar across shale counties, "boom towns," and the non-shale control superset during 2002-2017. Farm counts are less consistent, with shale areas having lower counts in the pre-2007 period and higher counts—particularly for shale counties

[38]Less than 4% of counties have suppressed hired worker data at any time in the 2002-2017 period. We impute hired worker counts to counties with suppressed data by comparing the sum of hired workers across counties to state-level hired worker counts available in higher tabulations without suppression; state discrepancies are shared out among counties with missing data based on each county's farm count as a share of total farms among counties in the state with missing hired worker data. Figure A7 shows data for all counties, including those with imputed hired worker data; in unreported results we find that dropping counties with imputed hired worker data has negligible effect on the results.

broadly—in the post-2007 period, suggesting that the shale boom might have had some positive effect on farm counts but not hired workers (i.e., average farm size would have seen a relative decline).[39] Understanding more about the effect of the shale boom on the agricultural economy is an important avenue for further research.

C.3 Decomposing employment growth

In section 4.2 we report a decomposition exercise assessing the contribution of establishment entry to overall employment growth. Here we describe the decomposition method. Let N_t be the number of establishments in year t, and let $\overline{emp}_t$ be average establishment size (employees per establishment) in year t. Consider the change in total employment between year 0 and year T, $N_T\overline{emp}_T - N_0\overline{emp}_0$. It can easily be shown that

$$N_T\overline{emp}_T - N_0\overline{emp}_0 = (N_T - N_0)\overline{emp}_0 + N_0(\overline{emp}_t - \overline{emp_0}) + (N_T - N_0)(\overline{emp}_T - \overline{emp}_0) \quad (15)$$

The first term on the right-hand side is the change in total employment accounted for by the change in the number of establishments (holding establishment size constant). The second term is the change in total employment accounted for by the change in the average establishment size (holding the number of establishments constant). The third term is a covariance term which, in practice, is relatively small. It is straightforward to calculate the share of total employment change that is accounted for by each of the components. Since the covariance term is small, for simplicity we distribute it proportionally among the other two terms.[40]

On Figure 4 in the main text, the solid line reports the total change in employment between 2006 and any given year ($N_T\overline{emp}_T - N_0\overline{emp}_0$). The dashed line reports the change in total employment in which establishment size is held constant ($(N_T - N_0)\overline{emp}_0$); this line indicates the portion of employment growth accounted for solely by the change in the number of establishments.

[39] In contrast with CBP and the LBD restriction to business establishments with paid employees, the Census of Agriculture includes farms both with and without hired labor. Roughly three-quarters of farms have no hired labor per the 2017 Census. For this reason, movements in farm counts can be less related to movements in farm employment. See Decker et al. (2021) for discussion of farm data and their relation to Census Bureau data sources.

[40] In all industries excluding oil and gas mining, the covariance term is less than 3%, while it is 24% in the oil and gas mining sector.

C.4 Effects on total employment by sector

Table A4 reports diff-in-diff results in which the dependent variable is (the log of) total employment; oil and gas mining is omitted, and results are reported for all plays, "boom towns," and each play separately (except Bakken and Anadarko, which are omitted for disclosure avoidance reasons). Table A5 reports these regresions by industry for all plays and for the "boom towns." We first study the oil and gas sector inclusive of both oil and gas mining (NAICS 211, 213) and related manufacturing (324, 325). Among all plays, this broad oil and gas sector saw average employment increased by almost 50% as a result of the shale boom; however, we find that this effect was driven entirely by the narrower oil and gas mining sector, which gained 70%, while the related manufacturing industries gains were not statistically significant. This latter finding reflects the fact that while significant downstream investments occurred in response to the shale boom, much of this investment was in areas with historical presence of these industries, not necessarily in new areas where extraction is now occurring,[41] therefore spurring significant investment in transportation infrastructure (Agerton and Upton, 2019; Agerton et al., 2020).

Employment outside of the oil and gas mining sector was also significantly affected, with impacts differing significantly across industries. For instance, construction, transportation and warehousing increased by 21.9%, while retail trade and leisure and hospitality experienced 3.6% and 7.3% increases, respectively with some other sectors such as utilities, professional business services and other services not statistically significantly impacted.

Additionally, we present results for each industry for only the "boom towns" sample. We find that similar industries are impacted for this sub-sample, but the magnitude of these effects is greater. Additionally, in contrast to the full results, we find a marginally statistically significant result (16%) for the oil and gas manufacturing sector; however, the magnitude of this effect is relatively small compared to the results for the oil and gas mining sector.

C.5 Alternative control groups

Our main results—and the causal interpretation thereof—depend on our propensity-matched control group. We first test the sensitivity of our results to alternative control groups by randomly choosing 20 control groups (rather than relying on our propensity score matching algorithm). The counties in these groups are drawn at random (with replacement) from

[41] Dismukes et al. (2019) estimates that over $110 billion in refining and chemical announcement occurred in Texas and Louisiana during the shale boom, but is mostly located near the Gulf Coast, not in the regions where the shale production actually occurred.

the our control group superset as described in the main text. We estimate our employment growth (by firm age) regressions with each of these 20 control groups; Table A10 shows the minimum, median, and maximum coefficients obtained from these control groups along with the propensity score match control group estimated treatment effects (i.e., repeated from Table A6).

The random control group exercises are generally supportive of our main results while pointing to the importance of our propensity score approach for generating causal inference. Column 7 of Table A10 reports coefficients for overall employment growth.

Broadly speaking, though, the random control group exercises support our main results and do not raise any concerns about our research design. The shale boom is plausibly exogenous to the patterns of business entry we study (particularly in industries outside oil and gas mining).

C.6 Placebo tests

We perform two placebo tests. We randomly assign observations to the control and treatment groups in two ways. First, we estimate our model only using the treated observations (i.e., counties in shale plays) but randomly assigning the observations to be "treatment" or "control". Second, we repeat this exercise using only the control observations (i.e., counties included in our propensity matched control group). Results of placebo tests for employment by shale play are presented in Table A11. None of the 14 coefficients in this table is statistically significant. Broadly speaking, our placebo tests are supportive of our identification strategy.

C.7 Accounting for shock magnitude

The size of the shale boom shock varied across plays and over time within plays. Our differences-in-differences estimate does not account for this heterogeneity. We can therefore gain more insights into our main annual employment growth results by allowing effects to vary by the size of the oil and gas boom. For simplicity, we do this by regressing employment growth components (for industries excluding oil and gas mining) on county-level oil and gas mining employment (in logs); while these regressions do not necessarily have a causal interpretation, they do directly relate the shale boom "shock" to its consequences for non-shale industries (alternatively, one may think of this exercise as a way of scaling treatment effects by treatment intensity). That is, we estimate:

$$g_{ct}^{j} = \alpha + \beta \ln emp_{ct}^{211,213} + \tau_c + \gamma_t + \varepsilon_{ct} \tag{16}$$

where $emp_{ct}^{211,213}$ is employment in NAICS 211 and 213, and fixed effects for county (τ_c) and year (γ_t) are included as before. We estimate this regression on the same sample as that used for Table A6; that is, we include both our treated counties and their matched control group. Table A9 ishows the results; since the independent variable is *the log of* oil and gas mining employment, we interpret the total effect (approximately) as follows: a 10% increase in county-level oil and gas mining employment is associated with an increase of *annual* overall employment growth (outside oil and gas mining) of 0.06 percentage points; new firms and greenfield establishments each account for more than a quarter of this overall effect.[42] We view this as a sizeable effect since oil and gas mining employment ultimately grew by roughly 200% in the shale areas.

Comparing these results with our main estimates, we find that the role of entry (both new firms and greenfields) is somewhat larger when we account for the magnitude of the shock to the oil and gas mining sector. When restricting the sample to "boom towns," this effect is even more pronounced; new firms and greenfields each account for more than 40% of the total employment growth effect. In unreported results, we also find that the entry margin is more important when scaling our differences-in-differences indicator by play-level annual oil and gas revenue (though comparisons with rig counts produce mixed results). The result that entry is more important when the independent variable (or treatment) is scaled by the size of the oil and gas mining shock is consistent with our theoretical discussion above. Business entry is highly (and disproportionately) responsive to aggregate shocks.

C.8 Job destruction and exit

A natural question is whether the strong entry responses we document occurred against a backdrop of higher job and business churning generally. In unreported results, we estimate our differences-in-differences specification with job destruction rates and establishment exit rates as dependent variables. Among all plays and among boom plays, we observe modestly negative but not statistically significant effects of the shale boom on both job destruction and establishment exit. In other words, the shale boom apparently did not raise overall job and business churn.

[42]These results are broadly similar if we lag oil and gas mining employment by one year.

C.9 Cumulative employment growth

In this section, we provide additional details on the cumulative growth results summarized in Section 6.3.

Table A12 reports effects on cumulative employment in the "boom towns," by year and relative to county-level employment in 2006, as described in Section 5.4 and equation (8); this table provides the estimates for Figure 5 discussed in the main text. Recall that for this purpose we discard the annually based definitions of firm and establishment entry used previously, instead focusing on establishments and firms created *during* versus *prior to* the shale boom broadly. Our discussion focuses on "boom towns," but results for "boom towns" are shown in Table A13 and Figure A9. Note that for any given year, the coefficients in columns 1, 2, and 3 sum to column 4. It is important to recall that these specifications include year fixed effects such that coefficients indicate employment relative to control group counties; roughly speaking this is still a difference in differences approach where we compare treatment county employment relative to 2006 to control county employment relative to 2006. The results can be interpreted as growth of group employment between 2006 and a given year, comparing treatment and control counties.

First, consider column 4 of Table A12 (and Table A13 for all plays), which reports the cumulative gain in total employment (in treatment versus control counties) for "boom towns." Prior to 2007, total employment is flat and close to zero (and not statistically significant), lending support to the assumption underlying our main difference-in-differences result that treatment and control counties have similar pre-treatment trends. After 2006, total employment rises, becoming statistically significant in 2008. By 2014, total employment in treatment counties has risen 36.5% in all areas since 2006 (relative to controls). The results still striking, albeit less extreme, when we examine all plays together; by 2014 there is a relative increase of 17.2% in total employment in the treated counties since 2006.

In column 1 we present results for establishments that were born prior to 2007 (that is, these establishments were incumbents when the shale boom began). We find a positive and significant effect of the shale boom for these establishments from 2008 onward. For example, in the year 2008, we find that employment among these pre-2007 establishment cohorts has risen 3.82% (1.29% among all plays, though not yet statistically signficant) relative to total employment in 2006. This effect peaks at 12.74% in 2013 (6.67% in "boom towns") before attenuating slightly to 11.09% in 2014 (5.47% in "boom towns"). If we divide the 2014 coefficient in column 1 by the 2014 coefficient in column 4, we find that these pre-2007 establishment cohorts account for about one-third of the total post-2006 rise in employment in shale areas (relative to control counties). The remaining two-thirds of the rise is therefore attributable to establishments born after 2006.

Focusing on "boom towns" (Table A12), in column 2 we present results for greenfield establishments, that is, those opened in 2007 or later by firms that existed prior to 2007. We see a positive and significant result beginning in 2009 (an increase in 1.75% of 2006 total employment) that strengthen annually to the end of the sample in 2014 (9.63% relative increase). This net job creation among new establishments of preexisting firms accounts for about a quarter of the cumulative gain in total employment as of 2014.

In column 3 we examine the effect of the shale boom on employment in new firms, that is, firms started in 2007 or later. Roughly speaking, these are firms that were created after the shale boom began. In these results, we again observe a positive, statistically significant treatment effect in 2009, consistent with the fact that new firms tend to start small, but by 2014 this group has a larger relative increase in employment (15.77%) than either of the other two groups. This net job creation among post-2006 firms accounts for 43% of total shale area employment growth relative to the counterfactual.

As noted above, new establishments (either born to preexisting firms or new firms) account for about two-thirds of the total employment gain. We graphically report these year-specific effects separately for each play on Figure A10. The results do vary notably by play; Eagle Ford and Permian Basin look similar to the overall results described above. The gas-heavy plays—Haynesville and Appalachia—show small overall employment effects and a less-consistent story about firm and establishment entry. Modest preexisting trend differences between treatment and control groups are sometimes evident in these areas, though the differences are rarely statistically significant. In short, however, the cumulative results suggest that areas in which the shale boom generated large economic expansions saw an important role for entry, with new firms ultimately accounting for the largest share of activity gains.

D Supplemental tables and figures

Table A1: Calibration details

Parameter	Description	Model 1	Model 2
β	Discount factor	0.96	0.96
α	Returns to scale	0.8	0.8
γ	Labor supply elasticity	2	2
ρ_z	Firm TFP persistence	0.55	0.55
σ_z	Firm TFP dispersion	0.22	0.22
μ_c	Fixed operating cost mean	-6.7	-6.7
σ_c	Fixed operating cost dispersion	0.9	0.9
c_e	Entry cost	e^{μ_c}	$3e^{\mu_c}$
ξ	Entrant signal shape		2.69

Table A2: Pre-boom county establishment counts by sector and play

	Min-ing	Util-ities	Con-struction	Man-ufacturing	Ret-ail	Transp-ortation & ware	Prof. Services & biz	Educ-ation & Health	Leis-ure & Hosp.	Other Services
Panel A: Treated and Control Areas										
All Shale Counties	21	6	147	66	234	46	199	160	137	161
Boom Towns	25	4	62	25	115	30	90	72	63	74
All Non-Shale	6	6	244	113	368	66	399	260	227	240
Potential Controls	4	6	287	137	423	77	460	299	262	271
Panel B: Major Shale Plays										
Anadarko	43	6	118	49	200	38	200	146	106	134
Appalachia	14	7	196	102	333	59	245	236	201	242
Bakken	8	3	31	9	55	14	30	25	36	36
Eagle Ford	12	5	63	24	124	63	72	66	64	69
Haynesville	27	11	98	52	212	38	148	129	93	131
Niobrara	25	6	304	91	318	60	456	225	201	207
Permian	27	4	45	20	89	19	60	53	48	59

Note: Average county-level establishment counts by play and NAICS sector for 2000-2006. Sectors are mining; utilities; construction; manufacturing; retail trade; transportation and warehousing; professional and business services; education and health; leisure and hospitality; and other services. Residual sectors omitted. Boom towns include counties in Anadarko, Bakken, Eagle Ford, and Permian Basin. Non-shale counties include all U.S. counties outside shale areas. Non-shale control set includes all counties except those in shale states and states adjacent to shale counties (see text). Source: County Business Patterns

Table A3: Pre-boom county employment by sector and play

	Min	Uti	Con	Man	Ret	Tran & ware	Prof & biz	Educ & heal	Leis & hosp	Oth svcs
Shale counties	320	160	1,180	2,790	3,060	700	2,630	3,720	2,250	1,050
Boom towns	330	80	480	800	1,370	360	1,030	1,520	1,040	490
Non-shale counties	130	210	2,200	4,780	4,970	1,320	6,500	5,960	4,100	1,790
Non-shale control set	80	230	2,560	5,840	5,760	1,530	7,470	6,430	4,830	1,980
Anadarko	440	130	900	1,870	2,450	600	2,610	2,880	1,940	990
Appalachia	310	250	1,450	4,790	4,460	1,060	3,570	6,060	3,160	1,510
Bakken	150	70	130	140	560	70	240	670	380	160
Eagle Ford	190	80	480	760	1,500	640	650	1,440	1,040	390
Haynesville	310	160	920	2,870	2,570	460	1,920	3,370	2,020	980
Niobrara	360	140	2,800	2,740	4,420	840	5,440	3,540	3,480	1,380
Permian Basin	400	50	390	500	1,050	210	640	1,140	810	400

Average county-level employment by play and NAICS sector for 2000-2006. Sectors are mining; utilities; construction; manufacturing; retail trade; transportation and warehousing; professional and business services; education and health; leisure and hospitality; and other services. Residual sectors omitted. Boom towns include counties in Anadarko, Bakken, Eagle Ford, and Permian Basin. Non-shale counties include all U.S. counties outside shale areas. Non-shale control set includes all counties except those in shale states and states adjacent to shale counties (see text). Source: County Business Patterns

Table A4: Effect of shale on (log) employment - All industries except oil and gas mining

	(1) All	(2) Boom Towns	(3) Appalachia	(4) Eagle Ford	(5) Haynesville	(6) Niobrara	(7) Permian
$\hat{\delta}$	0.069***	0.131***	0.014	0.198***	-0.008	0.0025	0.091***
	(0.012)	(0.023)	(0.012)	(0.054)	(-0.038)	(0.035)	(0.035)
N	9,420	3,780	3,780	690	750	1,110	1,620

Oil and gas mining sector (NAICS 211, 213) omitted. Dependent variable natural log of total employment excluding oil and gas mining employment in all regressions. County clustered standard errors shown. Treatment time period post 2007. Treated areas include all counties with shale oil and/or gas production as defined by EIA Drilling Productivity Reports. Control counties chosen using propensity score match from national sample in non-shale states. "Boom Town" is a combination of Permian, Anadarko, Eagle Ford, and Bakken plays. Bakken and Anadarko results are included in the "All" group but are not able to be reported individually due to data confidentially constraints. Parameters estimated with OLS. Source: Longitudinal Business Database

Table A5: Effect of shale on employment by industry

	(1) All Oil & Gas	(2) Upstream Oil & Gas	(3) Oil & Gas-Manufacturing	(4) Mining	(5) Utilities	(6) Const., Trans. & Warehousing
NAICS	211, 214, 324, 325	211, 213	324, 325	21	22	23,48,49
Panel A. All Areas						
$\hat{\delta}$	0.476***	0.701***	0.039	0.621***	0.015	0.219***
	(0.067)	(0.057)	(0.066)	(0.060)	(0.038)	(0.028)
N	9,420	9,420	9,420	9,420	9,420	9,420
Panel B. Boom Towns						
$\hat{\delta}$	0.593***	0.602***	0.160*	0.726***	0.085	0.414***
	(0.093)	(0.076)	(0.095)	(0.089)	(0.059)	(0.055)
N	3,780	3,780	3,780	3,780	3,780	3,780

	(7) Manufacturing	(8) Retail Trade	(9) Prof. Business Services	(10) Education & Health Services	(11) Leisure &, Hospitality	(12) Other Services
NAICS	31, 32, 33	44,45	54, 55, 56	61, 62	71, 72	81
Panel A continued. All Areas						
$\hat{\delta}$	0.107***	0.036***	0.012	-0.022	0.073***	0.023
	(0.035)	(0.013)	(0.035)	(0.021)	(0.020)	(0.017)
N	9,420	9,420	9,420	9,420	9,420	9,420
Panel B continued. Boom Towns						
$\hat{\delta}$	0.180***	0.085***	0.067	-0.031	0.110***	0.032
	(0.064)	(0.025)	(0.062)	(0.046)	(0.040)	(0.032)
N	3,780	3,780	3,780	3,780	3,780	3,780

Dependent variable natural log of total employment in all regressions. County clustered standard errors shown. Treatment time period post 2007. Treated areas include all counties with shale oil and/or gas production as defined by EIA Drilling Productivity Reports. Control counties chosen using propensity score match from national sample in non-shale states. "Boom Town" is a combination of Permian, Anadarko, Eagle Ford, and Bakken plays. Parameters estimated with OLS. Source: Longitudinal Business Database

Table A6: All plays: Effect of shale on employment growth - All industries except oil and gas mining

	(1) New Firms	(2) Young Firms	(3) New & Young Firms (1)+(2)	(4) Mature Firms	(5) Greenfield Estabs	(6) Incumbent Estabs	(7) Total (1)+(2)+(4) (1)+(5)+(6)
$\hat{\delta}$	0.321***	0.359***	0.680***	1.149***	0.340**	1.167***	1.829***
	(0.122)	(0.123)	(0.144)	(0.274)	(0.146)	(0.301)	(0.335)
N	9,420	9,420	9,420	9,420	9,420	9,420	9,420
Share of Total	17.6%	19.6%	37.2%	62.8%	18.6%	63.8%	100%

Oil and gas mining sector (NAICS 211, 213) omitted. Dependent variable growth component in all regressions. County clustered standard errors shown. Treatment time period post 2007. Treated areas include all counties with shale oil and/or gas production as defined by EIA Drilling Productivity Reports. Control counties chosen using propensity score match from national sample in non-shale states. "Boom Town" is a combination of Permian, Anadarko, Eagle Ford, and Bakken plays. Parameters estimated with OLS. New firm age (in years) =0, young =1-4, old = 5+. Columns 1+2=3, columns 1+2+4=7, and columns 1+5+6=7. Source: Longitudinal Business Database

Table A7: Effect of shale on employment growth - By industry

	(1) New Firms	(2) Young Firms	(3) New & Young Firms (1) + (2)	(4) Mature Firms	(5) Greenfield Estabs	(6) Incumbent Estabs	(7) Total (1)+(2)+(4) (1)+(5)+(6)
Panel A. All Areas							
Oil & Gas Mining (NAICS 211, 213)							
$\hat{\delta}$	7.84***	0.94	8.78***	10.37***	2.71	8.60**	19.15***
	(2.81)	(2.06)	(3.19)	(3.15)	(2.07)	(3.44)	(4.41)
N	5,133	5,133	5,133	5,133	5,133	5,133	5,133
Share of total	40.9%	4.9%	45.8%	54.2%	14.2%	44.9%	100%
Construction, Warehousing, & Transportation (NAICS 23, 48, 49)							
$\hat{\delta}$	0.97***	1.07***	2.04***	2.50***	0.22	3.36***	4.55***
	(0.37)	(0.32)	(0.41)	(0.64)	(0.29)	(0.69)	(0.79)
N	9,382	9,382	9,382	9,382	9,382	9,382	9,382
Share of total	21.3%	23.5%	44.8%	54.9%	4.8%	73.8%	100%
All Other Industries							
$\hat{\delta}$	0.18	0.28**	0.46***	0.91***	0.25*	0.94***	1.37***
	(0.13)	(0.12)	(0.14	(0.26)	(0.14)	(0.29)	(0.30)
N	9,420	9,420	9,420	9,420	9,420	9,420	9,420
Share of total	13.1%	20.4%	33.6%	66.4%	18.2%	68.6%	100%
Panel B. Boom Towns							
Oil & Gas Mining (NAICS 211, 213)							
$\hat{\delta}$	7.61***	2.45	10.06**	6.45*	3.99	4.90	16.51***
	(3.77)	(2.72)	(4.78)	(3.76)	(3.13)	(5.00)	(6.31)
N	2,125	2,125	2,125	2,125	2,125	2,125	2,125
Share of total	46.1%	14.8%	60.9%	39.1%	24.2%	29.7%	100%
Construction & Transportation (NAICS 23, 48, 49)							
$\hat{\delta}$	2.76***	2.25***	5.01***	5.28***	1.27**	6.26***	10.28***
	(0.85)	(0.69)	(0.93)	(1.25)	(0.59)	(1.33)	(1.61)
N	3,752	3,752	3,752	3,752	3,752	3,752	3,752
Share of total	26.8%	21.9%	48.7%	51.4%	12.4%	60.9%	100%
All Other Industries							
$\hat{\delta}$	0.86***	0.44	1.30***	1.22***	0.42	1.24**	2.52***
	(0.29)	(0.28)	(0.30)	(0.47)	(0.26)	(0.59)	(0.61)
N	3,780	3,780	3,780	3,780	3,780	3,780	3,780
Share of total	34.1%	17.5%	51.6%	48.4%	16.7%	49.2%	100%

County clustered standard errors shown. Treatment time period post 2007. Treated areas include all counties with shale oil and/or gas production as defined by EIA Drilling Productivity Reports. Control counties chosen using propensity score match from national sample in non-shale states. "Boom Town" is a combination of Permian, Anadarko, Eagle Ford, and Bakken plays. Parameters estimated with OLS. New firm age (in years) =0, young =1-4, old = 5+. Source: Longitudinal Business Database

Table A8: Effect of shale on employment growth - By play

	(1) New Firms	(2) Young Firms	(3) New & Young Firms (1) + (2)	(4) Mature Firms	(5) Greenfield Estabs	(6) Incumbent Estabs	(7) Total (1) + (2) + (4) (1) + (5) + (6)
Appalacia							
$\hat{\sigma}$	-0.325***	0.134	-0.192	-0.542	0.008	-0.416	-0.733**
	(0.116)	(0.089)	(0.126)	(0.274)	(0.169)	(0.321)	(0.194)
N	3,780	3,780	3,780	3,780	3,780	3,780	3,780
Eagle Ford							
$\hat{\sigma}$	0.873	0.116	0.989	1.479	1.723*	-0.128	2.468
	(0.624)	(0.580)	(0.603)	(1.06)	(0.837)	(1.169)	(1.499)
N	690	690	690	690	690	690	690
Haynesville							
$\hat{\sigma}$	0.325	0.376	0.701	1.794***	0.368	1.802**	2.495***
	(0.341)	(0.272)	(0.510)	(0.687)	(0.305)	(0.768)	(0.880)
N	750	750	750	750	750	750	750
Niobrara							
$\hat{\sigma}$	0.227	0.048	0.275	0.133	0.721	-0.541	0.408
	(0.426)	(0.307)	(0.437)	(1.071)	(0.462)	(1.108)	(1.238)
N	1,100	1,100	1,100	1,100	1,100	1,100	1,100
Permian							
$\hat{\sigma}$	0.699	0.722	1.421**	1.933**	0.899*	1.756	3.354***
	(0.485)	(0.530)	(0.570)	(0.939)	(0.479)	(1.126)	(1.196)
N	1,620	1,620	1,620	1,620	1,620	1,620	1,620

County clustered standard errors shown. Treatment time period post 2007. Treated areas include all counties with shale oil and/or gas production as defined by EIA Drilling Productivity Reports. Control counties chosen using propensity score match from national sample in non-shale states. Parameters estimated with OLS. New firm age (in years) =0, young =1-4, old = 5+. Due to data confidentially constraints, we are unable to report individual play results for Bakken and Anadarko. Source: Longitudinal Business Database

Table A9: Relationship of oil & gas mining employment with non-oil & gas mining employment growth

	(1) New Firms	(2) Young Firms	(3) New & Young Firms (1)+(2)	(4) Mature Firms	(5) Greenfield Estabs	(6) Incumbent Estabs	(7) Total (1)+(2)+(4) (1)+(5)+(6)
Panel A. All Areas							
$\ln emp^{211,213}$	0.142***	0.006	0.148	0.409***	0.163***	0.251	0.557***
	(0.050)	(0.118)	(0.136)	(0.111)	(0.053)	(0.181)	(0.206)
N	9,420	9,420	9,420	9,420	9,420	9,420	9,420
Share of Total	25.5%	1.1%	26.6%	73.4%	29.3%	45.1%	100%
Panel B. Boom Towns							
$\ln emp^{211,213}$	0.513***	-0.142	0.371	0.812***	0.475***	0.195	1.183*
	(0.156)	(0.384)	(0.443)	(0.284)	(0.128)	(0.552)	(0.625)
N	3,780	3,780	3,780	3,780	3,780	3,780	3,780
Share of Total	43.4%	12.0%	31.4%	68.6%	40.2%	16.5%	100%

Dependent variable growth component in all regressions. County clustered standard errors shown. Treatment time period post 2007. Treated areas include all counties with shale oil and/or gas production as defined by EIA Drilling Productivity Reports. Control counties chosen using propensity score match from national sample in non-shale states. "Boom Town" is a combination of Permian, Appalachian, Eagle Ford, and Bakken plays. Parameters estimated with OLS. New firm age (in years) =0, young =1-4, old = 5+. Columns 1+2=3, columns 1+2+4=7, and columns 1+5+6=7. Source: Longitudinal Business Database

Table A10: Employment growth effects with randomly selected control groups

	(1) New Firms	(2) Young Firms	(3) New & Young Firms (1) + (2)	(4) Mature Firms	(5) Greenfield Estabs	(6) Incumbent Estabs	(7) Total (1)+(2)+(4) (1)+(5)+(6)
Propensity Score Match Control Group							
$\hat{\delta}$	0.321	0.359	0.680	1.149	0.340	1.167	1.829
Random Control Groups							
$\hat{\delta}_{minimum}$	0.254	0.103	0.458	0.575	0.522	0.151	1.163
$\hat{\delta}_{median}$	0.355	0.256	0.597	0.769	0.641	0.338	1.339
$\hat{\delta}_{maximum}$	0.480	0.367	0.763	1.038	0.801	0.717	1.707

Treatment time period post 2007. Treated areas include all counties with shale oil and/or gas production as defined by EIA Drilling Productivity Reports. Control counties chosen using propensity score match from national sample in non-shale states. Data across all industries and shale plays is used. Propensity score match group coefficient estimates are from Table A6. Revenue is expressed in hundreds of millions of dollars and rig count is expressed in hundreds of rigs. Parameters estimated with OLS. New firm age (in years) =0, young =1-4, old = 5+. Columns 1+2=3, columns 1+2+4=7, and columns 1+5+6=7. Source: Longitudinal Business Database

Table A11: Placebo tests

	(1) New Firm	(2) Young Firm	(3) New & Young Firm (1) + (2)	(4) Old Firm	(5) New- Existing	(6) Old-Existing	(7) Total (1)+(2)+(4) (1)+(5)+(6)
Panel A: Treatment Placebo							
$\hat{\delta}$	-0.136	-0.067	-0.203	-0.398	0.241	-0.706	-0.601
	(0.197)	(0.214)	(0.233)	(0.422)	(0.232)	(0.497)	(0.534)
N	4,710	4,710	4,710	4,710	4,710	4,710	4,710
Panel B: Control Placebo							
$\hat{\delta}$	-0.055	-0.073	-0.128	-0.368	0.041	-0.483	-0.497
	(0.145)	(0.120)	(0.170)	(0.351)	(0.177)	(0.339)	(0.405)
N	4,710	4,710	4,710	4,710	4,710	4,710	4,710

Dependent variable natural log of total employment in all regressions. County clustered standard errors shown. Treatment time period post 2007. Treated areas include all counties with shale oil and/or gas production as defined by EIA Drilling Productivity Reports. Control counties chosen using propensity score match from national sample in non-shale states. Parameters estimated with OLS. New firm age (in years) =0, young =1-4, old = 5+. Source: Longidutinal Business Database

Table A12: Cumulative employment growth effects - Boom towns

	(1) Pre-2007 establishments *(Incumbent Estabs)*	(2) New establishments 2007 to pre-2007 firms *(Greenfield Estabs)*	(3) New establishments of firms born 2007 and later *(New Firms)*	(4) Total *(All)*
Boom Towns				
δ_{2000}	0.0206	0	0	0.0206
	(0.0240)			(0.0240)
δ_{2001}	0.0206	0	0	0.0206
	(0.0229)			(0.0229)
δ_{2002}	0.0268	0	0	0.0268
	(0.0193)			(0.0193)
δ_{2003}	0.0087	0	0	0.0087
	(0.0212)			(0.0212)
δ_{2004}	0.0010	0	0	0.010
	(0.0164)			(0.0164)
δ_{2005}	0.0030	0	0	0.0030
	(0.0119)			(0.0119)
δ_{2007}	0.0114	0.0038	-0.0058	0.0094
	(0.0106)	(0.0094)	(0.0047)	(0.0151)
δ_{2008}	0.0382***	0.0104	0.0369	0.0855**
	(0.0127)	(0.0072)	(0.0350)	(0.0382)
δ_{2009}	0.0832***	0.0175**	0.0136**	0.1143***
	(0.0161)	(0.0080)	(0.0057)	(0.0185)
δ_{2010}	0.0808***	0.0226**	0.0145*	0.1179***
	(0.0153)	(0.0107)	(0.0077)	(0.0186)
δ_{2011}	0.0992***	0.0413***	0.0294***	0.1699***
	(0.0181)	(0.0146)	(0.0093)	(0.0259)
δ_{2012}	0.1090***	0.0563***	0.0828***	0.2481***
	(0.0199)	(0.0161)	(0.0188)	(0.0396)
δ_{2013}	0.1274***	0.0821***	0.1178***	0.3272***
	(0.0212)	(0.0203)	(0.0257)	(0.0509)***
δ_{2014}	0.1109***	0.0963***	0.1577***	0.3649***
	(0.0221)	(0.0219)	(0.0328)	(0.0603)
N	3,780	3,780	3,780	3,780

County clustered standard errors shown. Base year 2006 and therefore not shown in table. Oil and gas mining (NAICS 211 and 213) omitted. Treated areas include all counties with shale oil and/or gas production as defined by EIA Drilling Productivity Reports. Control counties chosen using propensity score match from national sample in non-shale states. "Boom Town" is a combination of Permian, Anadarko, Eagle Ford, and Bakken plays. Parameters estimated with OLS. New firm age (in years) =0, young =1-4, old = 5+. Employment ratio is defined as the ratio of a given group's employment in a given year to the total county employment for that group in the base year of 2006. Pre-treatment period is 2000-2006 and post-treatment period is 2007-2014. Source: Longitudinal Business Database

Table A13: Cumulative employment growth effects - All plays

	(1) Pre-2007 establishments *(Incumbent Estabs)*	(2) New establishments 2007 to pre-2007 firms *(Greenfield Estabs)*	(3) New establishments of firms born 2007 and later *(New Firms)*	(4) Total *(All)*
All Areas				
δ_{2000}	0.0138	0	0	0.0138
	(0.0129)			(0.0129)
δ_{2001}	0.0143	0	0	0.0143
	(0.0119)			(0.0119)
δ_{2002}	0.0190*	0	0	0.0190*
	(0.0101)			(0.0101)
δ_{2003}	0.0086	0	0	0.0086
	(0.0103)			(0.0103)
δ_{2004}	0.0095	0	0	0.0095
	(0.0083)			(0.0083)
δ_{2005}	0.0007	0	0	0.0007
	(0.0062)			(0.0062)
δ_{2007}	0.039	0.0037	-0.0027	0.005
	(0.0062)	(0.0082)	(0.0023)	(0.0077)
δ_{2008}	0.0129	0.0082**	-0.0142	0.0174**
	(0.0009)	(0.0036)	(0.0142)	(0.0174)
δ_{2009}	0.0456***	0.0096**	-0.0038	0.0099***
	(0.0084)	(0.0053)	(0.0048)	(0.0106)
δ_{2010}	0.0467***	0.0135**	0.0088*	0.0691***
	(0.0084)	(0.0053)	(0.0048)	(0.0106)
δ_{2011}	0.0558***	0.0211***	0.0169***	0.0938***
	(0.0095)	(0.0068)	(0.0053)	(0.0134)
δ_{2012}	0.0604***	0.0234***	0.0430***	0.1268***
	(0.0105)	(0.0076)	(0.0088)	(0.0187)
δ_{2013}	0.0667***	0.0348***	0.0553***	0.1568***
	(0.0112)	(0.0093)	(0.0115)	(0.0235)***
δ_{2014}	0.0547***	0.0417***	0.0760***	0.1724***
	(0.0114)	(0.0100)	(0.0144)	(0.0273)
N	9,420	9,420	9,420	9,420

Oil and gas sector (NAICS 211, 213) omitted. County clustered standard errors shown. Base year 2006 and therefore not shown in table. Treated areas include all counties with shale oil and/or gas production as defined by EIA Drilling Productivity Reports. Oil and gas mining (NAICS 111 and 113) omitted. Control counties chosen using propensity score match from national sample in non-shale states. Parameters estimated with OLS. New firm age (in years) =0, young =1-4, old = 5+. Employment ratio is defined as the ratio of a given group's employment in a given year to the total county employment for that group in the base year of 2006. Pre-treatment period is 2000-2006 and post-treatment period is 2007-2014. Source: Longitudinal Business Database

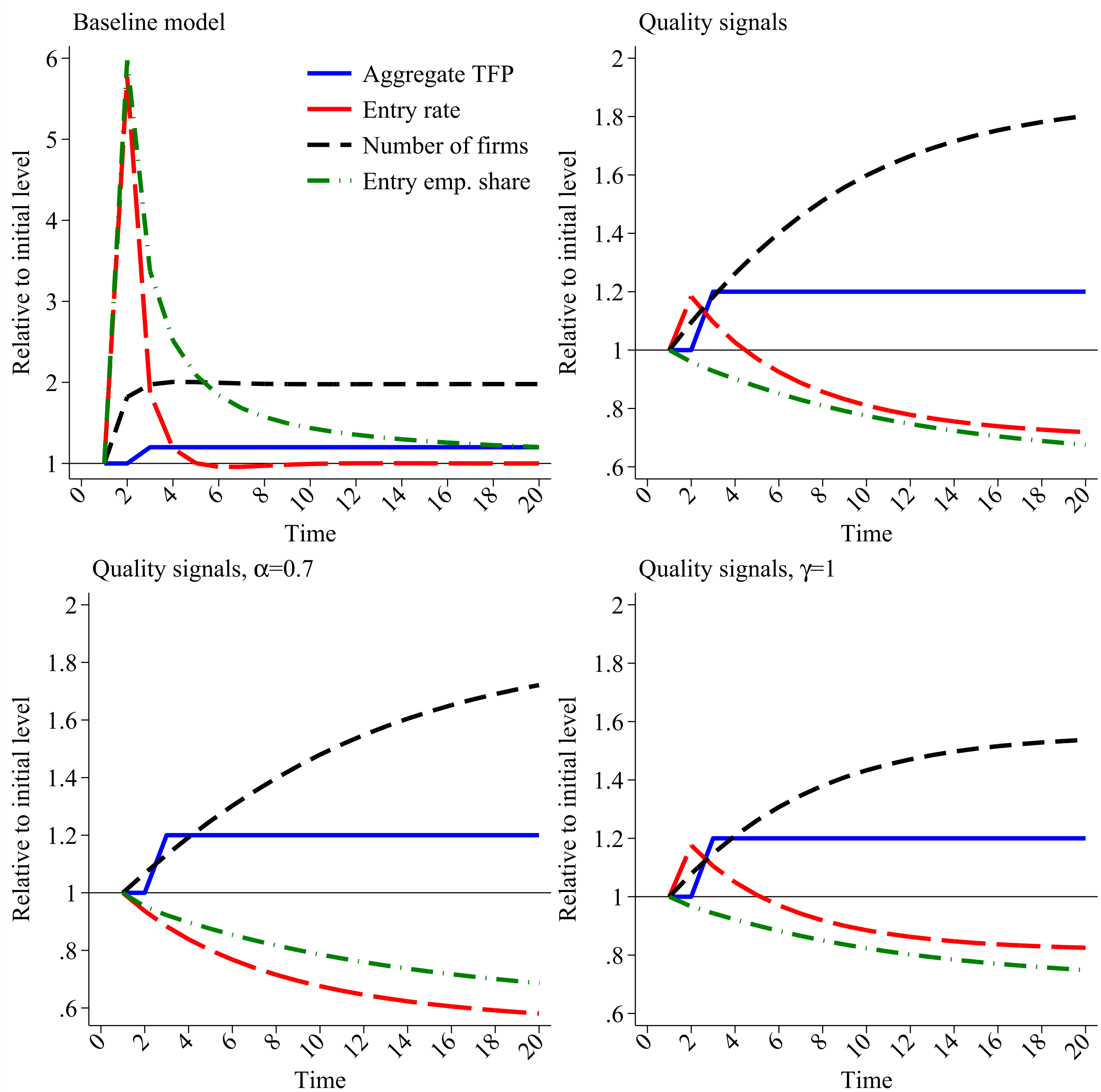


Figure A1: Model dynamics after aggregate productivity increase

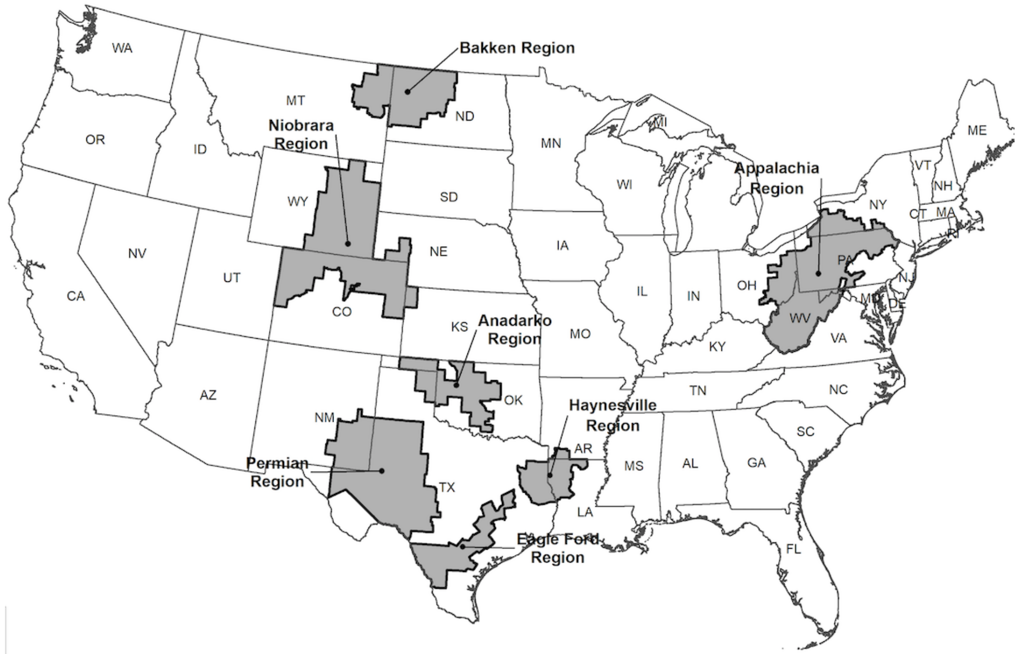
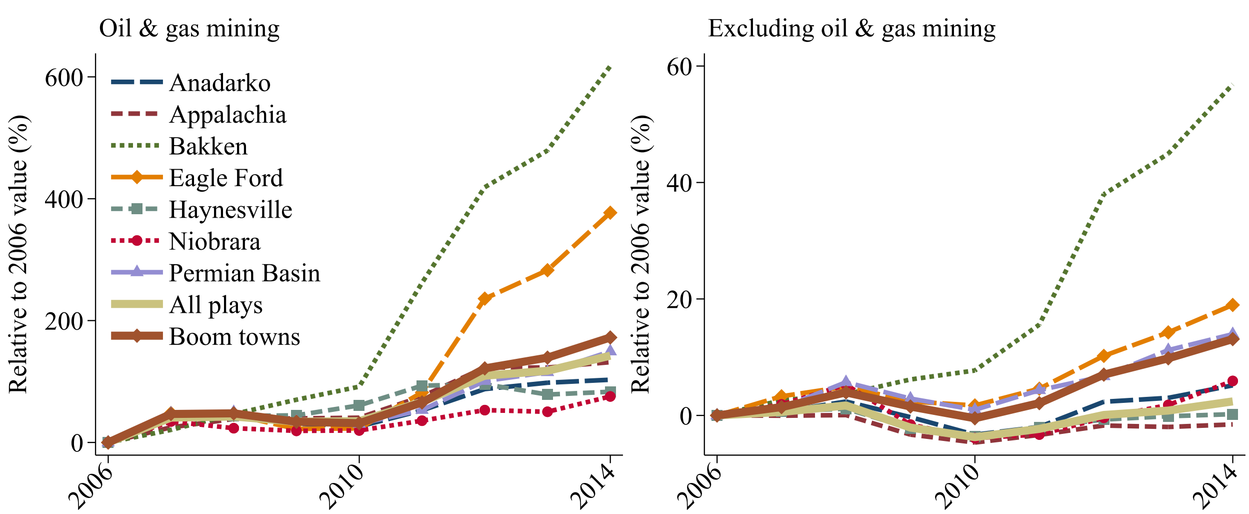


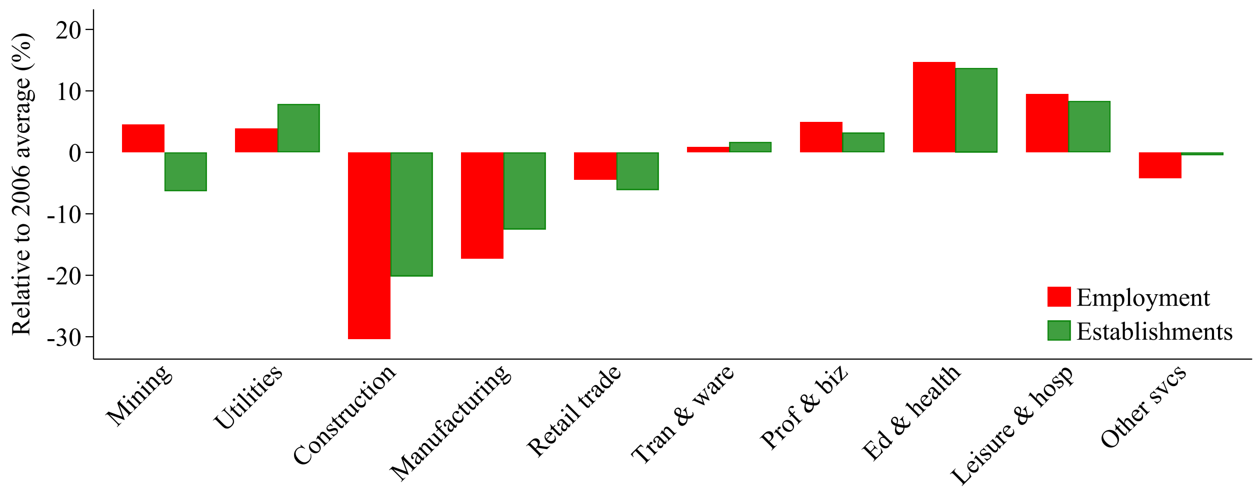
Figure A2: U.S. Shale Plays

Source: U.S. Energy Information Administration. Drilling Productivity Reports



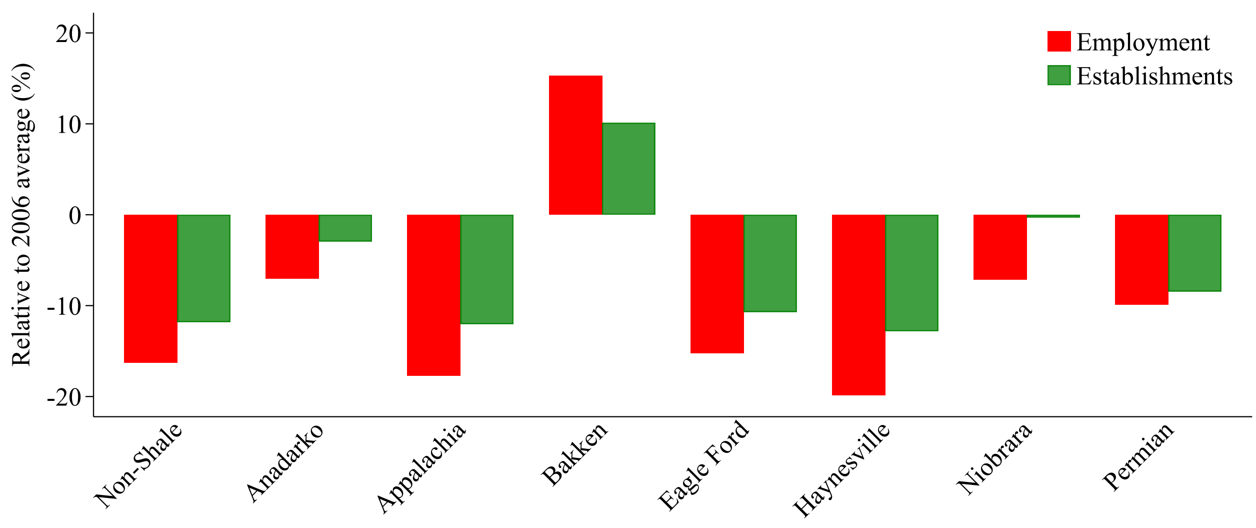
Note: Play-level employment relative to 2006. Boom towns are Anadarko, Bakken, Eagle Ford, and Permian Basin.
Source: County Business Patterns. Oil & gas mining is NAICS 211 and 213.

Figure A3: Employment gains by shale play



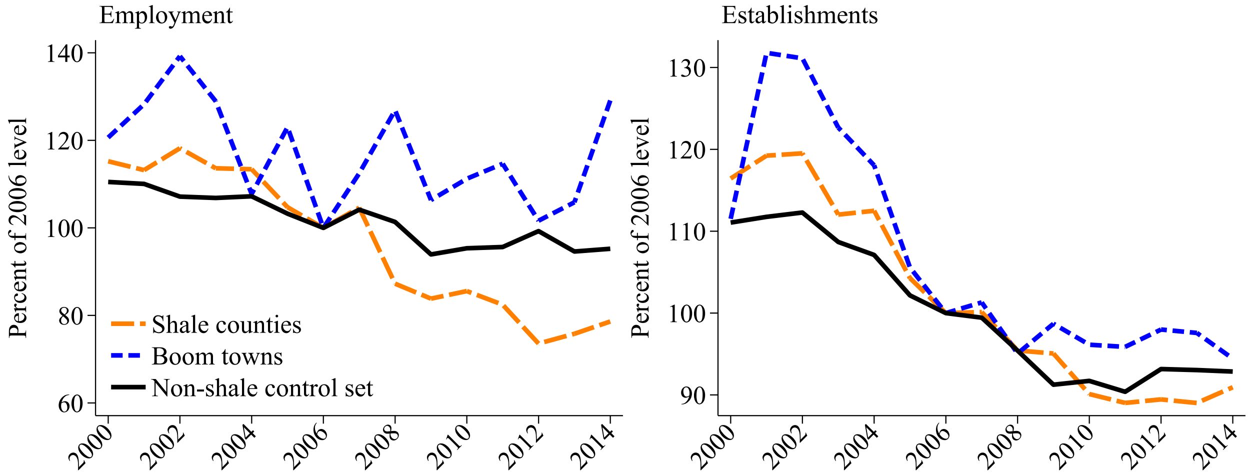
Note: 2006-2014 growth of average county-level employment and establishment counts. Excludes states with shale counties and states adjacent to shale counties. Residual industries omitted.
Source: County Business Patterns. NAICS sectors.

Figure A4: Control superset activity by sector, 2006-2014



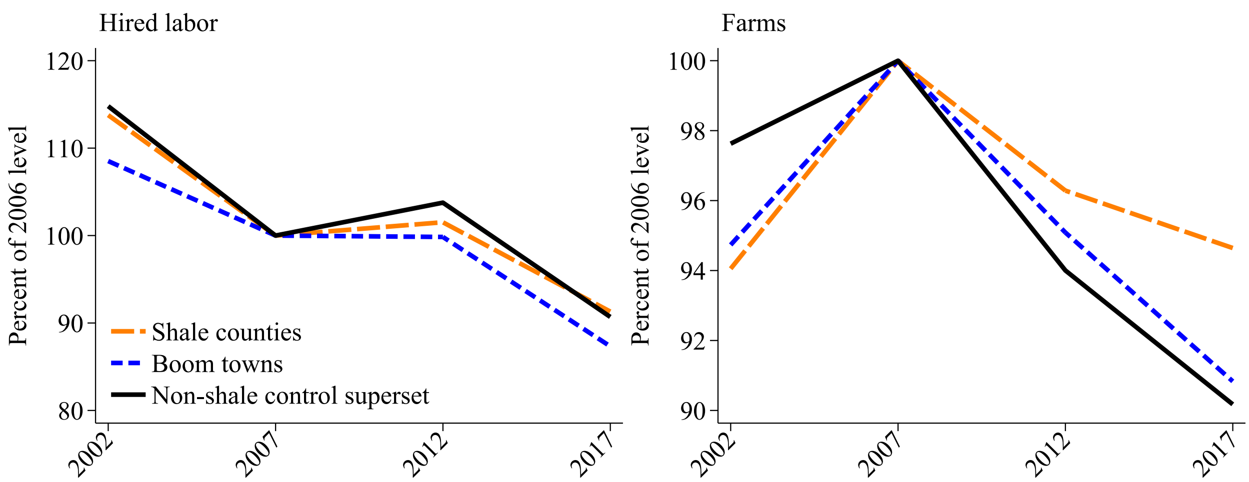
Note: 2006-2014 growth of average county-level employment and establishment counts.
Source: County Business Patterns. Manufacturing (NAICS 31-33) only.

Figure A5: Manufacturing activity by play, 2006-2014



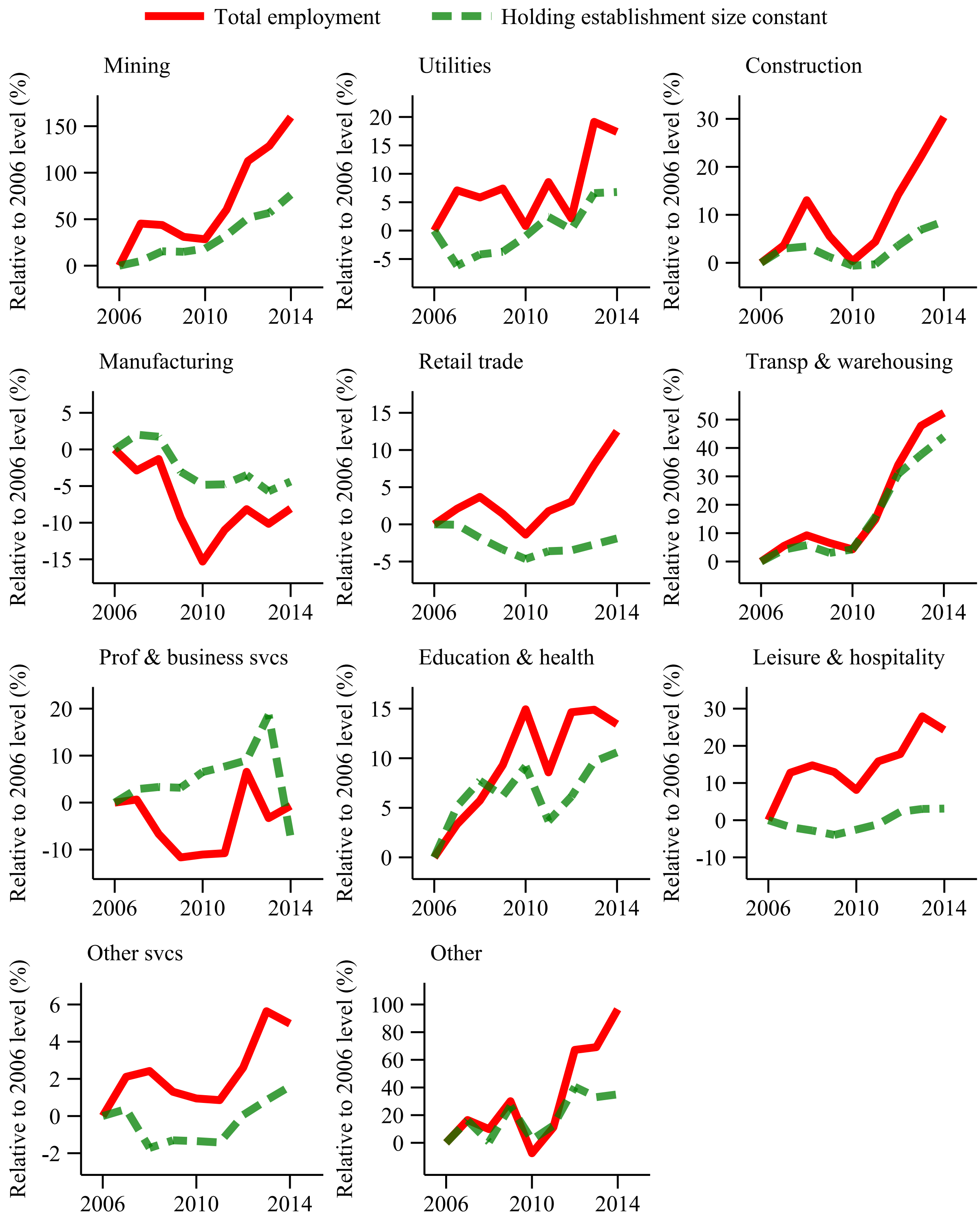
Note: Average county-level employment and establishment counts scaled by 2006 averages.
Boom towns are Anadarko, Bakken, Eagle Ford, and Permian Basin.
Source: County Business Patterns. NAICS 11 excl. 111 & 112. States that include or border shale counties omitted.

Figure A6: Non-farm agriculture activity, shale vs. non-shale counties



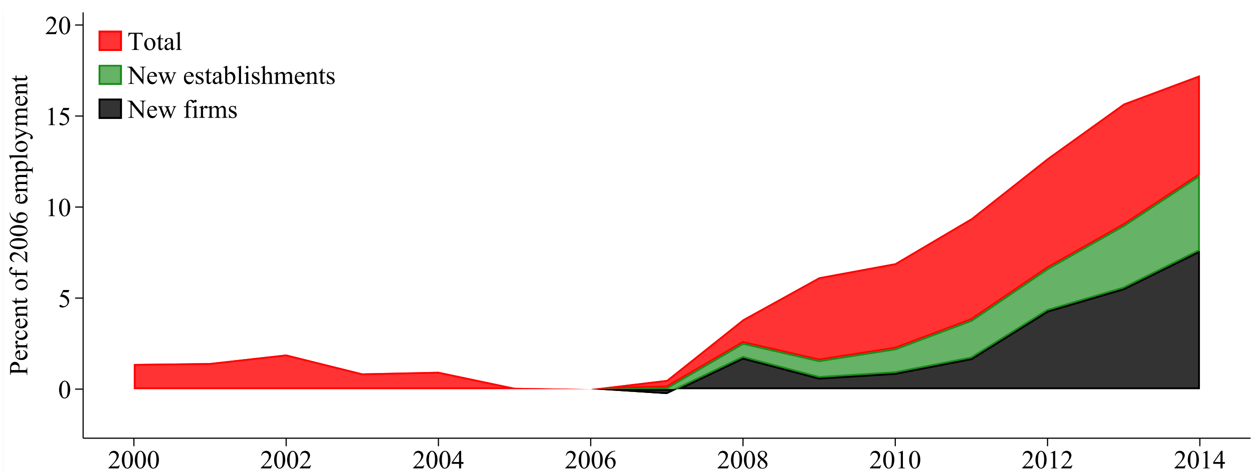
Note: Average county-level hired worker and farm counts scaled by 2007 averages.
Boom towns are Anadarko, Bakken, Eagle Ford, and Permian Basin.
Source: Census of Agriculture. NAICS 111 and 112. States that include or border shale counties omitted.

Figure A7: Farm activity, shale vs. non-shale counties



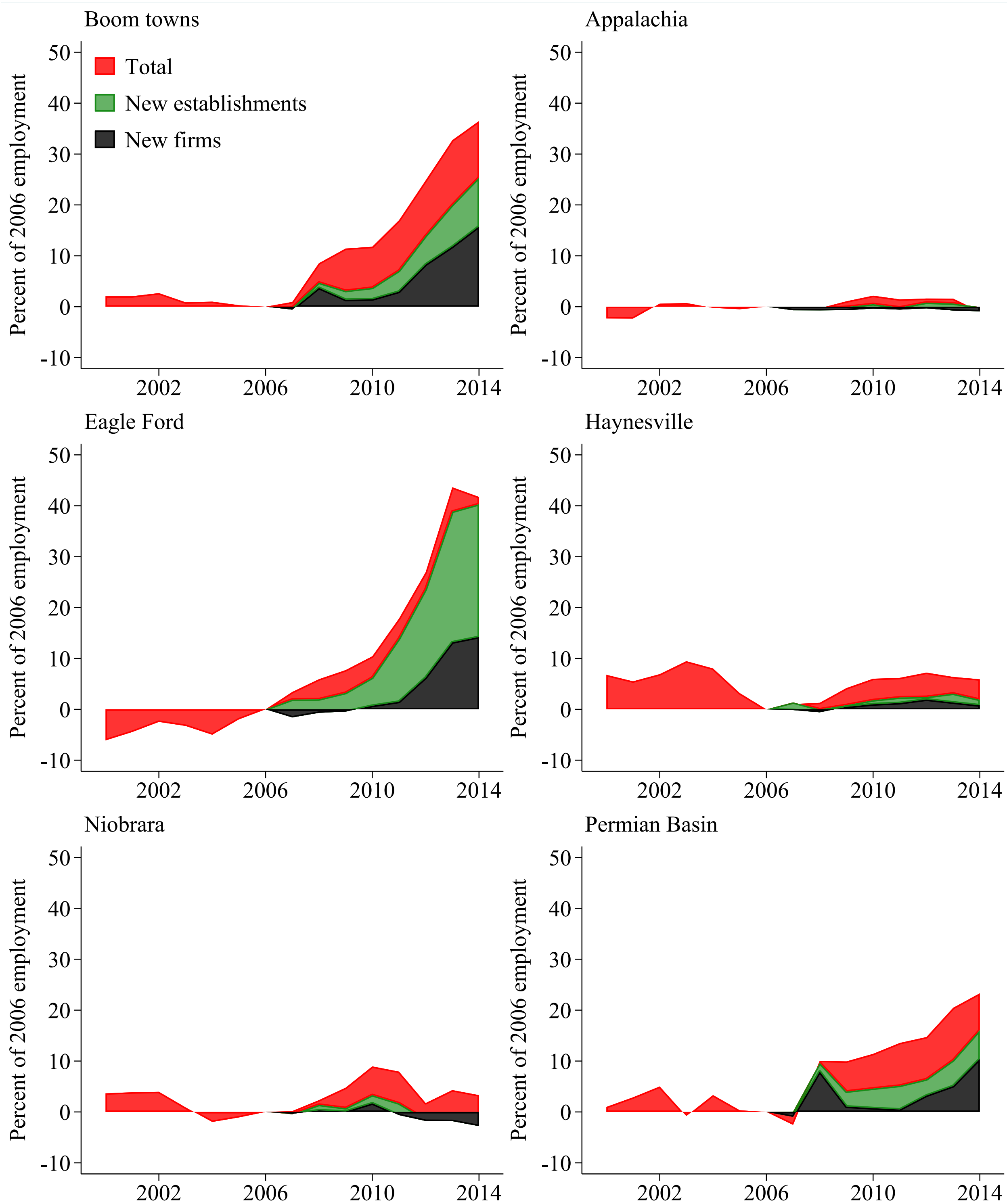
Note: Total employment relative to year-2006 level. Covariance term shared out proportionally.
Includes Anadarko, Bakken, Eagle Ford, and Permian Basin.
Source: County Business Patterns. NAICS sectors.

Figure A8: Margins of employment growth in boom towns by sector



Employment scaled by 2006 county employment. Regression compares treatment and control counties with year effects. New establishments are establishments born after 2006. New firms are firms born after 2006.
Author calculations from LBD. Oil and gas mining (NAICS 211, 213) excluded.
Total employment is statistically significant from 2008 on.

Figure A9: Employment treatment effects by year: All regions



Employment scaled by 2006 county employment. Regression compares treatment and control counties with year effects. New establishments are establishments born after 2006. New firms are firms born after 2006. 'Boom towns' are Permian, Anadarko, Eagle Ford, and Bakken plays. Author calculations from LBD. Oil and gas mining (NAICS 211, 213) excluded.

Figure A10: Employment treatment effects by year