# Entry, growth, and survival in the green industry

Dakshina G. De Silva and Timothy P. Hubbard and Robert P. McComb and Anita R. Schiller

Texas Tech University

31. October 2011

# Entry, Growth, and Survival in the Green Industry

Dakshina G. De Silva*
Timothy P. Hubbard†
Robert P. McComb†
Anita R. Schiller*

November 17, 2011

**Abstract**

Economists have, for some time, studied the factors that induce firm entry, lead to growth, and help firms succeed in various markets. Unfortunately, such patterns have not been considered for the so-called "green industries." Although policymakers might like to stimulate development of the green sectors in encouraging sustainable growth, one difficulty has been defining exactly what constitutes the green economy. We employ a recent, narrow definition proposed by the Bureau of Labor Statistics to investigate and identify important factors for the green industries within the State of Texas. We find some differences between the green industries and all other industries, but these effects are often small relative to other major explanatory factors like agglomeration. The definition also partitions the green industry into five subcategories and we leverage this feature to study the importance of these factors for the intra-green industries and to identify the comparative advantage each county has within the green economy.

**JEL Classification**: O44, Q56, R30.
**Keywords:** green industry, firm entry, employment growth, firm survival.

*De Silva (corresponding author, d.desilva@lancaster.ac.uk): Department of Economics, Lancaster University Management School, Lancaster University, Lancaster, LA1 4YX, UK; telephone +44/0 1524 594023; fax +44/0 1524 594244. Schiller (a.schiller@lancaster.ac.uk) is also a member of the Department of Economics at Lancaster University.

†Hubbard (timothy.hubbard@ttu.edu) and McComb (robert.mccomb@ttu.edu): Department of Economics, Texas Tech University, MS 41014, Lubbock, TX 79409-1014, USA.

# 1 Introduction

Economists have for many years worked to identify features of a national, regional, or local economy that attract investment and promote growth. Such information is useful to policy makers who are interested in attracting investment to boost employment and income. While researchers in industrial organization and regional economics have documented important factors for some time, we narrow our focus to firms that comprise the "green industry." Our objective in this paper is to analyze the factors that explain entry, growth, and exit patterns in the green industry to determine whether they differ from the brown industries. This insight is important given the increasing interest in sustainable development. For example, the Engrossed Second Substitute House Bill 2815 of 2008 aims to stimulate job creation in the "green economy" of Washington State.

One challenge that has precluded a thorough investigation of the green industry has been the lack of a clear definition of what exactly comprises this part of the economy. We take advantage of a recent definition proposed by the Green Jobs Initiative at the U.S. Bureau of Labor Statistics (BLS) based on six-digit level of the North American Industry Classification System (NAICS).[1] Broadly speaking, jobs are "green" if they are held at, or associated with, establishments that produce goods or provide services that benefit the environment or conserve resources (output-related green jobs) or if the jobs involve making environmentally-friendly production processes or focus on using fewer natural resources (process-related green jobs). We take this definition as given and sort all establishments in our data into one of two categories which, for simplicity, we refer to as the green and brown (otherwise known as non-green) industries.[2] While our primary interest is to evaluate whether the behavior of green firms is fundamentally different from that of brown firms, we also consider whether the "type" of green firm is important. Specifically, the BLS partitions the green industries into five subcategories: renewable energy (CAT1), energy efficiency (CAT2), pollution abatement and/or recycling (CAT3), natural resource conservation (CAT4), and environmental compliance, training, and awareness (CAT5).

We focus exclusively on the State of Texas during the period 2000-2006 and employ Quarterly Census of Employment and Wages (QCEW) data which allow us to observe key establishment-level variables at a reasonable level of industry disaggregation. Although we were forced to consider one state due to data limitations, the restriction is attractive in that environmental regulations are typically enacted at federal and state levels.[3] As such, any changes (observed or unobserved) in

[1]The website for the Green Jobs Initiative at the BLS, http://www.bls.gov/green/, describes information on the funding, development, and status of the program.

[2]This language and abstraction is analogous to researchers concerned with environmental economics who might distinguish brown, or polluting, firms from all other firms, which are then referred to as green. While this is common practice, given our research interest and the definition we employ, the partition works in the other direction for our research: we distinguish green firms from all other firms, which we refer to as brown firms.

[3]For example, Texas has a Renewable Portfolio Standard (RPS) concerning solar, wind, geothermal, hydroelectric, tidal energy, and biomass, which mandates "the construction of certain amounts of renewable energy and prompted the renewable energy industry to rapidly accelerate its production on Texas sites." (State Energy Conservation

state policies that occur during our sample obtain throughout the region of interest. This mitigates endogeneity issues that might arise with policy changes but also precludes our being able to say anything concerning the efficacy of such policies. Note, too, that since the BLS definition was proposed after the end of our period of analysis, there can be no endogeneity problems due to counties and municipalities attempting to attract the recently-defined green industries. It should be noted as well that Texas is a large and diverse economy and limiting the analysis to Texas is not, in fact, overly limiting. Indeed, as the second-largest state economy in the U.S. (after California), Texas ranked as the 15th largest economy in the world in a comparison of countries and states by gross domestic product, surpassing many notable national economies.[4]

Our research is complementary to that of researchers concerned with the effects of trade on the environment and the relationship between environmental policies and foreign direct investment (FDI). These researchers typically concentrate on identifying or testing whether brown (dirty) industries are attracted to certain locations based on pollution-based policies. Our focus is rather on identifying factors and patterns, if any, that appear important in attracting green firms and developing a sustainable economy.[5] Most notably, the pollution haven hypothesis is built on the endogenous response of a state (in intra-national models) or country (in international models) to attract dirty industries by reducing pollution standards. Empirical researchers often leverage differences in pollution policies (such as emissions tax rates) or regulation levels (such as attainment or non-attainment distinctions) to evaluate the effects on firm costs, entry, and growth. For example, Levinson [1996] found that interstate differences in environmental regulations did not affect the location choices of most manufacturing plants. However, List and Co [2000] found that state environmental regulations did alter multinational corporations' new plant location patterns. List [2001] considered how the number of FDIs in California were related to previous counts of FDI, market size, and land area. List found some evidence of agglomeration effects at the county level, which suggests that environmental regulatory policy may not be as important as benefits from agglomeration effects.

Many researchers have treated the subject of agglomeration economies (including knowledge spillovers) from various perspectives such as (Henderson [1986]), location choices of firms, (Rosenthal and Strange [2003] as well as Woodward et al. [2006]), firm exits (Staber [2001]), firm and industry growth (Glaeser et al. [1992], Henderson et al. [1995], as well as Combes [2000]), and labor

Office website: http://www.seco.cpa.state.tx.us/re_rps-portfolio.htm.) The Texas Legislature increased the state's renewable-energy goal in 2005. The RPS, however, does not favor one county (location) relative to another which is key for us given we use variation in county characteristics across Texas and over time to identify important features that explain green entry, grown, and survival.

[4]Gross domestic products for countries were obtained from The World Factbook produced by the Central Intelligence Agency while gross domestic products for states came from the Bureau of Economic Analysis.

[5]Other researchers who restrict attention to brown industries focus on quantifying the impact of environmental regulations on total factor productivity growth (for example, Barbera and McConnell [1990] considered five pollution-intensive industries), investigating the determinants of environmental innovation (for example, Brunnermeier and Cohen [2003] restricted attention to U.S. manufacturing industries), or other policy-driven effects.

productivity (Ciccone and Hall [1996]). Combes [2000] noted that a greater number of similar firms within a locality should increase the likelihood of complementary knowledge spillovers since there is greater likelihood of closer matches between firms. The consequence of this intra-industry knowledge spillover is, as Glaeser et al. [1992] pointed out, that "regionally specialized industries should grow faster because neighboring firms can learn from each other much better than geographically isolated firms." Ciccone and Hall [1996] also concluded that locally increasing returns to density explain more than half of the variation of labor productivity across states in the US. We include agglomeration effects (within a county as well as in contiguous counties) and knowledge spillover effects (either through firm agglomeration effects or by university and junior college research funding) in our firm entry, employment growth, and survival models.

Researchers try to link potential consequences of a policy using observed outcomes and investigate whether the effects are significant (and substantial). Similarly, Devereux et al. [2007] considered whether government subsidies (discretionary grants) affected where domestic and multinational firms located new plants. They found firms to be quite insensitive to government policies (consistent with much of the conclusion reached by those studying the pollution haven hypothesis) and more attracted to areas offering, for example, co-location benefits. This suggests that the intrinsic features of locations, whether exogenously fixed (such as resources) or endogenously determined (such as market structure and firm agglomeration), is more important. Thus, if the regulatory environment is not a factor, can it be that green and brown industries are attracted to locations for similar reasons? Some evidence for potential differences was provided by Eichholtz et al. [2010] who found that buildings with green ratings (characterized by Leadership in Energy and Environmental Design (LEED) or Energy Star certification) garnered significantly higher commercial office rents, suggesting that such classification has economic value.

We find that the green and brown industries appear to respond to similar non-policy factors and, in general, entry and growth patterns are quite similar across industries. Agglomeration effects are most important in explaining firm entry and employment growth. Although agglomeration effects are not fundamentally different in attracting green industries, they help employment growth in green industries more than in brown industries. However, employment growth in green industries is positively correlated with income, which is not true for brown employment. Concerning survival, green firms are more likely to exit compared to brown firms, although this effect is weakened if the green firm has previous experience. Moreover, agglomeration effects are important in terms of reducing establishment mortality. In addition to investigating the entry, growth, and survival patterns for the intra-green subcategories, we identify which counties within Texas have comparative advantages in each of these subcategories.

Our paper is structured as follows: in section 2, we describe how we identified green industries using the BLS definition while, in section 3, we discuss our data and present some general patterns we observe. We investigate these patterns in further detail by estimating various empirical models of firm entry, employment growth, and firm survival in section 4. In section 5, we identify comparative advantages for each county within Texas based on the green subcategories and, in section 6, we summarize and conclude.

## 2 Defining the Green Industry

On July 15, 2009, in order to measure green jobs accurately, the BLS created a discussion draft for the Workforce Information Council. The main objective of this was "to produce objective and reliable information on the number of green jobs, how that number changes over time, and the characteristics of these jobs and the workers in them." In addition to partitioning the number of jobs by industry that are associated with green good and services (GGS) production, the BLS was interested in estimating the occupational employment and wages for establishments identified as producing these GGS. The BLS hoped that this information would provide policymakers and the public with a better understanding of green jobs so that they could make informed decisions.

After surveying government, academic, and business studies, the BLS consulted with federal agencies, state labor market offices, and industry groups to better understand activities related to preserving or restoring the environment through efforts such as the production of renewable energy, improving energy efficiency, preventing and/or cleaning up pollution, and conserving natural resources.[6] In particular, green jobs are either

(a) Jobs in businesses that produce goods or provide services that benefit the environment or conserve natural resources.

(b) Jobs in which workers' duties involve making their establishment's production processes more environmentally friendly or use fewer natural resources.

After realizing that there was no formal understanding for what constituted the green industry, the BLS developed its own objective and measurable definition. Specifically, 333 six-digit industries from the 2007 NAICS were identified as green.[7] The BLS also partitioned this general definition of the green industry into subcategories. Specifically, of the 333 six-digit NAICS industries, 58 were categorized as relating to energy from renewable sources, 140 as relating to energy efficiency, 124 as relating to pollution reduction and removal, greenhouse gas reduction, and recycling and reuse, 75 as relating to natural resource conservation, and 45 as relating to environmental compliance, education and training, and public awareness. Summing the number of members of each of these subcategories suggests 442 industries, not 333. These values are reconciled because many industries

[6]Our discussion summarizes relevant material from the Green Jobs Initiative website provided in footnote 1 as well as the Federal Register Notices published on this site.

[7]We employ the final definition which was announced in volume 75, number 182 of the Federal Register Notice. The official industry list can be accessed at http://www.bls.gov/green/final_green_def_8242010_pub.pdf

are classified into more than one of the subcategories. For example, NAICS code 221330 is specified as "steam and air-conditioning supply" and is classified in the green subcategories concerning energy from renewable sources, energy efficiency, as well as pollution reduction and removal, greenhouse gas reduction, and recycling and reuse. For each of the 333 industries, the BLS provided examples suggesting why the industry was included. For example, the industry might be specified as producing or certifying organic foods, as relating to LEED, as producing, repairing, or certifying Energy Star products, as providing mass transit systems, etc. We take the BLS classification as given and we employ the BLS NAICS-based definition in our analysis below to identify green and brown establishments.

## 3 Data

We obtained data for this study from two primary sources. First, we use firm-level data for the State of Texas from the Quarterly Census of Employment and Wages (QCEW) provided by the Texas Workforce Commission. The data concern monthly employment and quarterly total wages reported by every establishment in the state as required under the Texas unemployment insurance program. Each record includes the specific location (address) of the establishment, the business liability start-up date (the date from which unemployment insurance liability begins), and the relevant six-digit NAICS code which is of particular importance for our work. Note, too, that separate establishments (branches or franchises) of the same firm are distinguished and reported in unique records. This panel data set is comprised of observations from the first quarter of 2000 (2000Q1) through the fourth quarter of 2006 (2006Q4), constituting 28 data periods.[8] Each record also includes each establishments's unique Employer Identification Number (EIN). Therefore, the appearance of a new EIN is used to define market entry and the disappearance of an EIN is treated as an exit.[9] One difficulty we faced, given our data come from the pre-2007 era, concerned the mapping between the 2002 NAICS codes which characterize the QCEW data and the 2007 NAICS codes which were used by the BLS to define the green categories. To identify the green categories in the 2002 NAICS classification we used the concordances provided by the U.S. Census Bureau.[10] We discuss details of this mapping in the appendix, in which we also provide a table which describes the variables we use in our analysis.

In table 1, we compare the firm (EIN establishment) and employment distribution across the

[8]The data were provided under an agreement of confidentiality and disclosure of the data is subject to certain restrictions.

[9]Following Dunne et al. [2005], some EINs appear in a given quarter but are associated with previous EINs. We do not treat such observations as new entrants given we can identify the prior EIN. The change in EIN may have occurred because the establishment changed hands, a partnership was broken up, or for any number of reasons. Unfortunately, we are not provided with any justification for the new EIN, but our data do allow us to recover the relationship to a previous EIN. In our firm survival analysis, we control for these "firms with past experience."

[10]The Census Bureau maintains concordances at http://www.census.gov/eos/www/naics/concordances/concordances.html.

Table 1: Distribution of Brown and Green Activity across Sectors

| Sector | Brown Firms | Green Firms | Brown Employ. | Green Employ. |
|---|---|---|---|---|
| Ag., Forestry, Fishing & Hunting | 0.19 | 7.47 | 0.07 | 2.68 |
| Mining | 1.74 | 0.00 | 2.25 | 0.00 |
| Utilities | 0.18 | 1.70 | 0.31 | 3.51 |
| Construction | 0.48 | 33.82 | 0.96 | 24.99 |
| Manufacturing | 4.86 | 5.13 | 8.23 | 19.37 |
| Wholesale Trade | 9.33 | 0.44 | 5.77 | 0.39 |
| Retail Trade | 17.87 | 1.05 | 15.76 | 0.48 |
| Transportation & Warehousing | 3.95 | 0.36 | 5.51 | 1.55 |
| Information | 1.43 | 3.39 | 2.50 | 5.36 |
| Finance and Insurance | 7.97 | 0.19 | 5.75 | 0.07 |
| Real Estate, Rental, & Leasing | 5.90 | 0.00 | 2.23 | 0.00 |
| Prof., Scien., & Technical Services | 5.03 | 27.03 | 1.96 | 15.50 |
| Management of Companies & Enterprises | 0.13 | 0.73 | 0.06 | 1.65 |
| Admin., Support, Waste & Remediation Serv. | 4.87 | 4.35 | 8.09 | 3.25 |
| Educational Services | 1.12 | 1.05 | 6.71 | 14.18 |
| Health Care & Social Assistance | 12.41 | 0.00 | 13.51 | 0.00 |
| Arts, Entertainment, & Recreation | 1.28 | 0.40 | 1.01 | 1.04 |
| Accommodation & Food Services | 9.61 | 0.00 | 10.36 | 0.00 |
| Other Services (except Public Admin.) | 9.98 | 10.32 | 2.31 | 3.43 |
| Public Administration | 1.65 | 2.57 | 6.65 | 2.58 |
| Total | 100.00 | 100.00 | 100.00 | 100.00 |

two-digit 2002 NAICS sectors conditional on being classified as a part of the brown or green industries. The numbers in the table correspond to the percent of green and brown firms, respectively, or the share of green and brown employment, averaged over the sample periods, that is attributed to each of the sectors listed. As such, the columns sum to 100 percent. Before interpreting the table, note that, in total, 23.32 percent of Texas firms are part of the green industry representing 18.01 percent of total employment. There are some stark differences in how economic activity is distributed across the green and brown industries. Agricultural firms account for a much larger share of the green industry as many are concerned with the production of or services related to organic produce and meat although, as a share of green employment, the sector is far less substantial. A third of green firms are considered part of the construction sector, while a quarter of green employees are construction-related. The high share of green employment in construction is primarily due to LEED policies and Energy Star certification. For example, many six-digit 2002 NAICS sectors produce or install LEED-eligible materials or concern installation of efficient environmental control systems. Likewise, the high share of firms and employment in the professional, scientific, and technical services is due to land surveying, architectural services, and energy- or resource-efficient design services, again often relating to LEED. Although the shares of green and brown firms in the manufacturing and educational services sectors are comparable, the shares of green employment in these sectors far outranks that of brown employment, suggesting these green firms are larger than their brown counterparts within these two-digit sectors, at least on average. This is not surprising for educational services as the six-digit codes defined as “junior colleges” as well as “colleges and universities” are both considered part of the green industry since they provide training for green jobs. Unfortunately while the BLS justified their classifications of which sectors belonged in the green industry, no rationale was provided for why some sectors were not part of the green industry and so we cannot offer any formal insight into why sectors concerned with real estate, health care, and accommodation and food services have no green firms (and, hence, employment). We remain agnostic concerning the classification of the green and brown industries and take the BLS definition as given.

We observe all QCEW variables across the 28 quarters constituting our data sample, in each of the 1,299 NAICS codes (that existed in Texas for at least one quarter and survived our data “cleaning” described in the appendix) for each of Texas’ 254 counties providing us with 9,238,488 raw observations. We aggregate entry, exit, and employment variables given the green industry definition (and our green-brown partition), yielding 14,224 ($= 28 \times 254 \times 2$) observations, two per county in each quarter, for our interindustry analysis and 35,560 ($= 28 \times 254 \times 5$) observations, five per county in each quarter, for our intra-green industry analysis. In table 2, we provide some summary information concerning incumbency, entry, exit, industry size, and quarterly wages in the green and brown industries as well as within the green industry.[11] The mean values reported weight all counties and quarters equally and so they should be interpreted as representing the average county at a given time in our sample.

Based on the number of incumbent firms, the brown industry is over three times larger than the green industry, although using employment numbers suggests the brown industry is over four and a half times larger, implying brown firms are typically larger than green firms. We condition exit statistics based on two types of firms: those that enter sometime during our sample which we also observe exit (new firm exits) and those that were already in the market when our sample began, but which we observe exit during our sample (old firm exits).[12] For establishments that we observe enter our sample, the number of entrants and exits are quite close to each other, suggesting the two rates are comparable which is consistent with the findings of Dunne et al. [1988]. Firms that were already in the sample when our data period began are typically less likely to exit the market,

[11] We measure firm exit by the presence (absence) of a firm’s account number in consecutive quarters. For example, if a firm is present in quarters 1 and 2, but absent in quarter 3, we infer that the firm exited the industry during quarter 2. Because of this approach, we have to throw out the final quarter for our exit analysis as we cannot decipher which firms remained in the industry after our sample period terminated.

[12] This partition is necessary to prevent aggregate exit rates from being nearly twice entry rates. To see this, the average number of total exits can be obtained by summing the number of entrant and incumbent exits in a given row of table 2. This is deceiving given the firms considered in the old firm exists do not affect the entry rates.

Table 2: Green vs. Brown Industry and Green Subcategory Summary Statistics

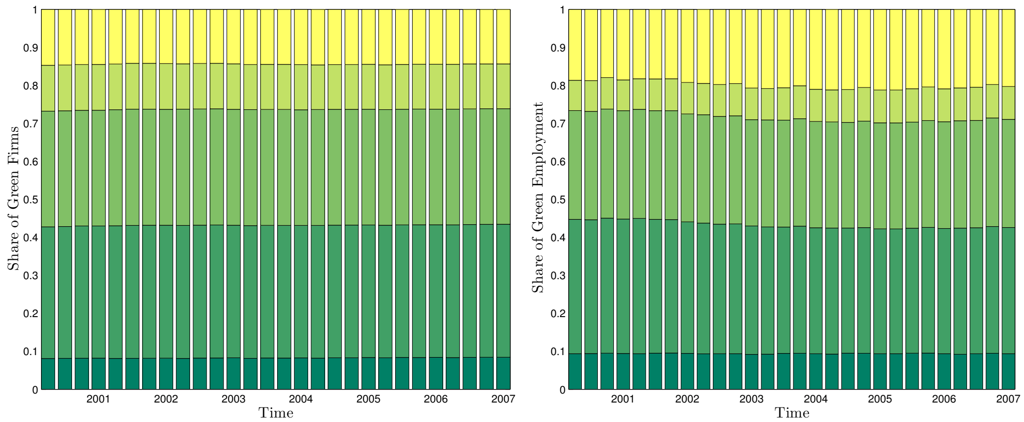
| | Number Incumbents | Number of Entrants | Number of New Firm Exits* | Number of Old Firm Exits* | Number of Employees | Quarterly Wages |
|---|---|---|---|---|---|---|
| Green | 423.65 | 11.36 | 11.00 | 10.56 | 9446.40 | 7634.60 |
| | (1574.61) | (46.59) | (48.28) | (40.04) | (45786.85) | (2897.18) |
| Brown | 1394.11 | 36.59 | 35.62 | 32.49 | 43001.47 | 6821.17 |
| | (5591.37) | (163.58) | (170.04) | (132.70) | (208570.20) | (1670.26) |
| CAT1 | 55.88 | 1.67 | 1.93 | 1.43 | 1360.24 | 9497.85 |
| | (233.25) | (7.45) | (8.24) | (6.09) | (7755.70) | (4662.23) |
| CAT2 | 235.58 | 7.52 | 7.60 | 5.76 | 4906.88 | 7766.29 |
| | (977.46) | (32.08) | (33.36) | (24.79) | (27123.73) | (3689.31) |
| CAT3 | 205.90 | 6.44 | 6.45 | 4.98 | 4094.40 | 7676.20 |
| | (825.98) | (27.40) | (28.55) | (20.69) | (21927.56) | (3086.04) |
| CAT4 | 81.01 | 1.88 | 1.80 | 2.16 | 1226.81 | 6852.95 |
| | (200.99) | (6.57) | (6.45) | (5.38) | (5430.48) | (3061.03) |
| CAT5 | 98.54 | 1.95 | 2.39 | 2.62 | 2885.99 | 6977.42 |
| | (405.21) | (9.62) | (11.60) | (11.03) | (13254.26) | (3691.74) |

Standard deviations are in parentheses below each mean value.

* New (old) firm exits are exits by firms that we observed enter (were already in) our data sample which began in the first quarter of 2000.

on average, although this is not true for the subcategories CAT4 and CAT5. Comparing wages suggests why policymakers might want to attract green firms: the average quarterly wage in the green industry is over $800 more than in the brown industry. The intra-green industry subcategories show the average wage can be over $2676 higher (CAT1) compared with the brown industry mean wage. CAT2 and CAT3 are the largest subcategories, on average, in terms of both the number of firms, entry into these industries, and the number of employees. They also provide above-average paying jobs, even compared with the green industry as a whole. The only subcategory that comes close to matching the average firm size (computed by dividing the average number of employees in a category by the average number of incumbent firms) of the brown firms is CAT5, perhaps because it contains education centers which are typically large.

In figures 1(a) and 1(b), we depict how the composition of the green industry in TX has changed over our data sample based on the number of firms and employees. The bar for a given quarter represents the share of green firms (employees) considered to be part of each green subcategory, with CAT1 being the bottom bar, then CAT2, CAT3, and CAT4, so that CAT5 is the top bar. The figures show remarkable stability within the green industry and for all subcategories over the data period. Both figures show that CAT2 and CAT3 are not only bigger on average (as was shown in table 2), but that they are the largest green subcategories in Texas for every quarter we observe. Because the figures only convey how the composition of the green industry evolved over time, they mask the overall trend in the number of firms and employees. In figures 2(a) and



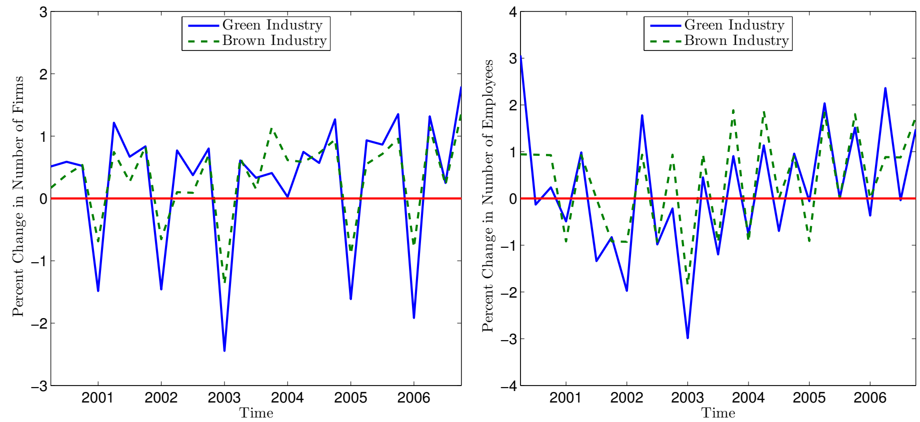

(a) Firm Composition

(b) Employment Composition

Figure 1: Composition of Green Industry in TX over Data Period

2(b), we depict the percentage change in the number of firms as well as the number of employees, respectively, across the quarters we observe in our data. While the changes in the number of firms and employees fluctuate between relative gains and losses, the two series seem to follow the same trend. These trends for the green subcategories follow a similar pattern and we do not present them formally as the figures became too cluttered. We found, at the aggregate state level, the most volatile subcategory in terms of the variance of the percentage change in employment was CAT5, followed by CAT4 and CAT1, respectively. However, CAT5 was the most stable in that the variance in the percentage change in the number of firms was lowest across our data periods, while CAT1 and CAT4 had the highest and second highest variances in their percentage change in the number of firms.

While tables 1 and 2 capture some summary statistics concerning the green and brown industries and figures 1(a)–2(b) depict how the industries have changed over the sample, neither of these provide any insight into the spatial distribution of the industries across Texas—something that might be important and help inform our empirical models. To better capture how the green and brown industries are distributed across the state we constructed maps of the mean number of green and brown incumbant firms in each county over the 28 data periods and of the mean number of green and brown entrants for a given county over the 28 periods. Perhaps not surprising, the major metropolitan areas as such as Dallas, Houston, San Antonio, and Austin, have the highest concentration of the green and brown industries both in terms of the number of incumbents and the number of entering establishments, suggesting that firms look to locate in either high population areas, or areas where other firms are already established. Note, too, that a spatial pattern developed around these MSAs: the MSA counties have the highest concentration of firms, neighboring counties




(a) Trend in Firms (b) Trend in Employment

Figure 2: Percentage Changes in Green & Brown Industries in TX over Data Period

appear to have slightly less activity, and rural counties have the lowest absolute number of firms.[13] Rather than present these four maps, we present in figure 3 a map depicting for each county the mean green intensity—a relative measure computed as the number of green firms in a county over the number of brown firms in a county. For reference, we have labeled the 25 metropolitan statistical areas (MSAs) defined by the U.S. Census Bureau which are based off population.[14] Surprisingly, some counties have more green firms than brown firms, as indicated by a green intensity greater than one. Moreover, a map considering an employment-based measure of the green intensity has the same qualitative features.

These various slices of our data, as summarized by the tables, figures, and maps above suggest patterns we investigate in more detail in the next section. However, they also inform us of features that might be important for our empirical models: first, accounting for both population density and agglomeration effects will be important; second, there may be spillover effects between neighboring counties; third, the stylized facts we hope to characterize may be inherently different for MSA counties than for the State of Texas as a whole. To account for other factors that might be important in explaining firm entry, employment growth, or firm survival, we complement the QCEW data described above with data from other sources. Specifically, county-level data as such as population density, were collected from the U.S. Census Bureau's Annual Population Estimates. We also calculated the average quarterly county income by taking the average wages paid in the county for

[13] To save space, we have not included these four maps (an incumbent and entry map each for the green and brown industries), but they are available from the corresponding author upon request.

[14] We list only the largest city in our labels. For example, the largest MSA in Texas is Dallas–Fort Worth–Arlington which we have labeled "Dallas" to avoid overlapping labels. This seems consistent with the way the Census Bureau identifies the areas which is motivated by size considerations and not an alphabetical one; for example, "San Antonio" corresponds to the MSA "San Antonio–New Braunfels."

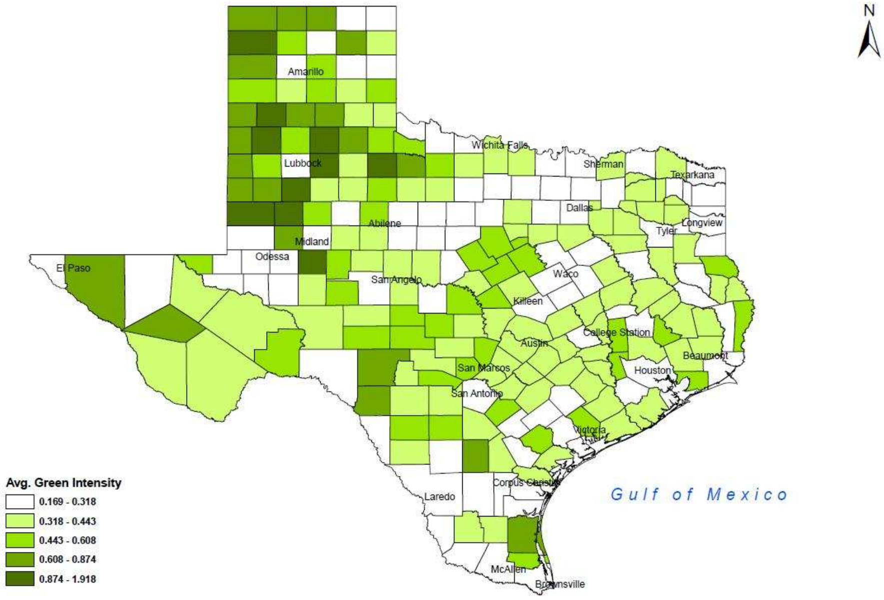



Figure 3: Average Distribution of Green Firm Intensity

all establishments reported in our QCEW data. Income seems to be not only a natural attractor of firm investment, but may be particularly important for green investment and/or employment: if green goods are normal goods then higher incomes would induce increased demand for green goods and services. In our empirical models, we investigate whether income has a different effect on the green industry compared with the brown industry.

In addition to the agglomeration and income effects we've formalized, we investigated whether the presence (and magnitude) of a research center affects the green industry. Having research universities provides access to expert consultants and allows for, often specialized, funding sources which may play a significant role in attracting green industries. For example, Abramovsky et al. [2007] found evidence that business-sector research and development activity is often located near university research departments. In order to capture such effects we identified the local presence of a four year university, junior college, or a research institution. Data on annual university research and development (R&D) expenditures were obtained from the National Science Foundation.

To account for factor costs, as in Bresnahan and Reiss [1991], we used the yearly median undeveloped land price in each of 33 land market regions in Texas for the counties comprising the regions as reported by the Texas A&M Real Estate Center. Given this variable does not change across quarters and is common to groups of counties, we also included the county-specific property tax rate. Lastly, as Woodward et al. [2006] suggested, cultural and natural amenities are important to industrial attraction and skilled workforce retention. As in De Silva and McComb [forthcoming], we focused on the share of county employment in local cultural and recreational amenities as being potentially important. This measure was intended to capture the influence of the locality's urban amenities on its attractiveness. While natural amenities may be valued, urban amenities are both more immediate and relevant to day-to-day life for full-time employed individuals. These activities also reflect the scope of the locality's amenities for business travelers as well as informal business and social interaction.

# 4 Empirical Models & Estimation

While the summary information we've presented thus far suggested some similarities between the green and brown industries, we would like to both identify factors that are important to understanding the green industry (and how it differs from the brown industry) and to control for things that may have changed over our data period and across counties which may have muddled the effects we are trying to uncover. In this section, we describe formal econometric models we estimated to better understand if there are inherent differences between the green and brown industries as well as within the subcategories of the green industry. In the spirit of Dunne et al. [1988, 1989], our interest is in firm entry, employment growth, and survival. We partition this section into three subsections corresponding to each of these topics.

## 4.1 Entry

Firm entry helps promote competition and improve efficiency. An immediate impact of new firms on a local economy is that they allow for job creation and attracting green investment is particularly attractive to policymakers concerned with sustainable development. In this subsection, we describe our investigation into the factors that attract firms in the green industry with particular emphasis on whether these factors have different effects as compared with the brown industry. To the greatest possible extent, we used a canonical set of explanatory variables that has appeared in the industrial organization and regional economics literatures. There has been interest in both fields in spatial effects, expressed through agglomeration economies, that attract firms and exhibit a self-reinforcing tendency to grow. We captured agglomeration effects by computing the number of firms already in an industry, within a given county at a given time. Agglomeration can be particularly important in location decisions when proximity to market is not a dominant factor. Thus, where localization leads to pooling of green labor, facilitation of communication among suppliers, access to intermediate inputs, and technological spillovers, an existing industry concentration increases the attractiveness of a locality for an establishment surveying areas in which to locate. Any clustering of green industries may also be the result of a deeper regional environmental consciousness insofar as it reflects social receptivity and interest in green activities. While we cannot account explicitly for local attitudes toward "going green" given our data, we included county fixed effects in our model to help capture unobserved, county-specific effects that were constant throughout our data period. Likewise, if these attitudes (or other factors, such as the overall health of the U.S. and Texas economies) are changing over time but are common to all counties, then we can capture these unobserved effects by including quarter-specific fixed effects in our model.

To consider factors that affect entry (firm investment), we constructed counts of the number of firms within each industry (or subcategory) that entered each county in a given quarter and considered a Poisson model with fixed effects to help control for unobserved heterogeneity. Researchers who adopt a standard Poisson model assume the dependent variable $y$ is independently distributed and the distribution of $y$ is a Poisson distribution. A consequence of adopting this assumption is that it imposes equality between the mean and variance of the dependent variable, conditional on explanatory variables; i.e., $\mathbb{E}(y|\mathbf{x}) = \mathbb{V}(y|\mathbf{x})$. However, a specification test for overdispersion in our data rejects the standard Poisson assumption. One alternative would be to use a negative binomial model where the assumed distribution for the dependent variable exhibits overdispersion. However, the negative binomial model is only consistent if the conditional distribution of the dependent variable is in fact negative binomial. Moreover, Guimarães [2008] showed that that the conditional maximum likelihood estimator of the negative binomial with fixed effects does not necessarily remove the individual fixed effects in count panel data. Specifically, this happens only if the number of groups is at least 1000 with more than 20 periods per group.

In contrast, the Poisson model is very robust in this aspect: regardless of whether the Poisson distribution holds, a consistent and asymptotically normal estimator can be obtained via quasi-maximum likelihood (QML) estimation. This is, perhaps, not surprising when comparing this with the ordinary least squares estimator which shares the consistency and asymptotic normality properties, regardless of whether the error terms are normally distributed. For the Poisson case, Gourieroux et al. [1984] showed that a consistent and asymptotically normal estimator can be obtained without specifying the probability density function of disturbances representing specification error in the parameter of the Poisson distribution. Likewise, Wooldridge [1999] showed that the fixed effects Poisson estimator is consistent and asymptotically normal as long as

$$\mathbb{E}(y_{ijt}|\alpha_j, \mathbf{x}_{ijt}; \boldsymbol{\beta}) = \alpha_j \exp(\mathbf{x}'_{ijt}\boldsymbol{\beta})$$

where, in our case, $y_{ijt}$ represents the number of entering firms in industry $i$, in county $j$, at time $t$, and $\alpha_j$ is fixed effect for county $j$. Furthermore, Wooldridge derived a robust covariance matrix for the Poisson QML estimator with conditional fixed effects. Thus, given our interest in the effects of the explanatory variables on the mean response, we estimated a Poisson model by QML.[15]

While the estimated coefficients obtained from Poisson QML estimation are identical to Poisson regressions with fixed effects, the standard errors must be adjusted for overdispersion.[16] In table 3,

[15] Cameron and Trivedi [2005] described and compared methods for estimating cross-sectional count data models in Chapter 20 of their book.

[16] Simcoe [2007] provided an implementation of the Poisson QML model with conditional fixed effects suggested by

we report results from four models estimated via Poisson QML in which robust standard errors were clustered by counties. We constructed these models to uncover any differences in the entry behavior between green and brown firms as well as within the green industry by comparing subcategory entry patterns. To investigate potential differences we considered two samples: a full sample involving all counties in the State of Texas and a restricted sample involving only counties that are considered to be part of a Texas MSA. This restriction is motivated by an observation that most entry occurred around population centers. In all models, we included as conditioning variables measures of county income, agglomeration within a county—computed as the number of like firms already present in the county of a certain type (green, brown, or belonging to a specific category), agglomeration in neighboring (contiguous) counties, university and junior college funding for each county, the county unemployment rate, the population density of the county, the undeveloped land price for the market region to which the county belongs, as well as county and time fixed effects to account for county-level and time-specific unobserved heterogeneity.[17] In table A.1 in the appendix we describe formally how these variables were constructed.

The four models in table 3 are distinguished by the comparison (green versus brown industries or a intra-green industry comparison) or by the sample (full sample versus the counties comprising the Texas MSAs). An initial comparison of the entry behavior between green to brown firms might consider whether, conditional on all other covariates, the average number of entrants (captured by a green dummy variable) suggests differences. Both samples indicate there is no significant difference in the likelihood of green and brown firms entering a county, all else equal. Note too, that income effects are not important in attracting new firms in any of our models. The most important factors in explaining firm entry are agglomeration and population density. For example, the incidence rate ratio calculated using the agglomeration variable in model (1) is 2.12, suggesting that a one percent increase in agglomeration attracts about two new entrants per county per quarter, although the effect is not significantly different for the green industry. This is consistent with List [2001] (as well as others cited in our introduction) who found that agglomeration is the primary factor in driving firm entry. Overall, the Poisson QML estimation results suggest no difference between green and brown entry patterns.[18]

If the sample of observations is restricted to only green entrants, we can investigate whether entry patterns differ across subcategories. Here our benchmark is CAT5, which we omitted. Entry is more common in CAT1, CAT2, and CAT3, all else equal, as shown by the subcategory dummy

Wooldridge [1999].

[17]In explaining entry, all right-hand side variables were lagged by a quarter to reflect the county environment at the time each new firm entered. As such, we had to throw out the first period of data as lagged variables are not available for the entry counts in the first data period.

[18]We also considered estimation of the models based on only the green industries and only the brown industries (which have the advantage of essentially interacting the green dummy variable with each of the covariates) which would be analogous to the approach taken by List and Co [2000]. Results are very similar but harder to interpret given the sample changes across model runs—this is particularly relevant for comparing the subcategory estimates to those of the green industry as a whole. These results are available from the corresponding author upon request.

Table 3: Poisson QML Estimation Results Concerning the Number of Entrants

| Variable | All Counties | | | | MSA Counties | | | |
|---|---|---|---|---|---|---|---|---|
| | Gr. vs. Br. (1) | | Gr. Subcats. (2) | | Gr. vs. Br. (3) | | Gr. Subcats. (4) | |
| Green industry | -.451 | (.564) | | | -.022 | (.658) | | |
| CAT1 | | | 2.129 | (1.581) | | | 2.974* | (1.720) |
| CAT2 | | | 1.937* | (1.011) | | | 2.076** | (1.055) |
| CAT3 | | | 2.157** | (.939) | | | 2.275** | (1.027) |
| CAT4 | | | 2.111 | (1.708) | | | 1.339 | (1.785) |
| Log (income)$_{j,t-1}$ | -.252 | (.155) | -.080 | (.219) | -.200 | (.219) | -.009 | (.2597) |
| Log (income)$_{j,t-1}$×Green industry | .042 | (.071) | | | .003 | (.082) | | |
| Log (income)$_{j,t-1}$×CAT1 | | | -.174 | (.192) | | | -.290 | (.208) |
| Log (income)$_{j,t-1}$×CAT2 | | | -.088 | (.124) | | | -.105 | (.133) |
| Log (income)$_{j,t-1}$×CAT3 | | | -.137 | (.115) | | | -.149 | (.129) |
| Log (income)$_{j,t-1}$×CAT4 | | | -.190 | (.208) | | | -.117 | (.217) |
| Log (agglomeration)$_{i,j,t-1}$ | .753** | (.044) | .734** | (.050) | .730** | (.053) | .743** | (.067) |
| Log (agglomeration)$_{i,j,t-1}$×Green industry | -.006 | (.010) | | | -.016 | (.009) | | |
| Log (agglomeration)$_{i,j,t-1}$×CAT1 | | | -.102* | (.044) | | | -.106** | (.047) |
| Log (agglomeration)$_{i,j,t-1}$×CAT2 | | | -.078* | (.036) | | | -.077* | (.043) |
| Log (agglomeration)$_{i,j,t-1}$×CAT3 | | | -.058* | (.035) | | | -.062 | (.044) |
| Log (agglomeration)$_{i,j,t-1}$×CAT4 | | | -.093 | (.058) | | | -.049 | (.071) |
| Log (agglomeration in neighbors)$_{i,j,t-1}$ | .024 | (.029) | .075 | (.052) | .090 | (.057) | -.011 | (.064) |
| Log (agglomeration in neighbors)$_{i,j,t-1}$×Green industry | -.013 | (.012) | | | -.007 | (.013) | | |
| Log (agglomeration in neighbors)$_{i,j,t-1}$×CAT1 | | | .044 | (.037) | | | .067* | (.039) |
| Log (agglomeration in neighbors)$_{i,j,t-1}$×CAT2 | | | .005 | (.036) | | | .014 | (.043) |
| Log (agglomeration in neighbors)$_{i,j,t-1}$×CAT3 | | | .008 | (.035) | | | .016 | (.042) |
| Log (agglomeration in neighbors)$_{i,j,t-1}$×CAT4 | | | .051 | (.044) | | | .022 | (.051) |
| Log (college funds)$_{j,t-1}$ | .003 | (.002) | .004 | (.003) | .003 | (.002) | .004 | (.003) |
| Log (junior college funds)$_{j,t-1}$ | .000 | (.001) | -.000 | (.001) | .000 | (.001) | .000 | (.001) |
| Unemployment rate$_{j,t-1}$ | -.005 | (.009) | -.037** | (.014) | -.021** | (.010) | -.062** | (.014) |
| Log (population density)$_{j,t-1}$ | .416** | (.127) | .311* | (.172) | .365** | (.173) | .333 | (.210) |
| Property tax rate $_{j,t-1}$ | -.056 | (.172) | .100 | (.221) | .063 | (.239) | -.043 | (.289) |
| Log (undeveloped land price)$_{j,t-1}$ | .029 | (.025) | .014 | (.050) | .004 | (.030) | -.007 | (.051) |
| Amenities employment ratio$_{j,t-1}$ | .663 | (.633) | .886 | (.826) | 1.134 | (1.128) | 1.237 | (1.362) |
| County effects | Yes | | Yes | | Yes | | Yes | |
| Time effects | Yes | | Yes | | Yes | | Yes | |
| Number Obs. | 13716 | | 34290 | | 4266 | | 10665 | |
| Wald $\chi^2$ | 123347.99 | | 66454.89 | | 11196.90 | | 60670.22 | |

** Denotes statistical significance at the 5% level and * denotes statistical significance at the 10% level.

Robust standard errors are in parentheses.

variables. These subcategories have slightly lower agglomeration effects, although the total effects from agglomeration are still positive and significant; however, like the green-brown comparison, the agglomeration effects do not seem to cross county borders as agglomeration in contiguous counties is not important in explaining firm entry. We also do not find evidence of knowledge spillovers (beyond that captured by the agglomeration variables) as measured by our research funding variables.

## 4.2 Employment Growth

To evaluate whether there are differences in the growth of the green and brown industries (or between subcategories within the green industry), we considered a simple regression model to try and explain the percentage change in industry (subcategory) employment of the counties. Specifically, in comparing the green and brown industries, we estimated the following model by ordinary least squares:

$$\log\left(E_{ijt}+1\right)=\mathbf{x}_{jt}'\boldsymbol{\beta}+\varepsilon_{ijt}$$

where the error term consists of independent shocks

$$\varepsilon_{ijt}=\alpha_i+\gamma_j+\theta_t+u_{ijt}.$$

We account for industry-specific unobserved heterogeneity $\alpha_i$, county-specific unobserved heterogeneity $\gamma_j$, and time-specific unobserved heterogeneity $\theta_t$, by including industry, county, and time fixed effects. The $u_{ijt}$ represents idiosyncratic error accounting for unobserved factors that change over time and across industries as well as counties. In our notation, $E_{ijt}$ denotes total employment in industry $i$, in county $j$, during period $t$. For our intra-green industry subcategory analysis we use the same model but $E_{ijt}$ is employment in subcategory $i$ for a county in a given quarter.

In table 4, we present results from estimation of these linear regression models. Because the full and restricted sample results are at least qualitatively similar, we interpret the results in a way that applies to both columns (1) and (3) of the table. The employment growth within a county is higher for brown industries, all else equal—the BLS Green coefficient suggests that the percentage change in employment in green industries will be 5.275% less than that of the brown counterpart. Higher incomes lead to a negative effect (a decrease) in brown employment, while the trend is reversed for green employment, perhaps offering some (indirect) justification for assumption that green goods and services are normal goods—as is often maintained in the environmental economics literature. Agglomeration effects help increase growth in industry employment and the effect on green industries is slightly higher (and significant). These results are consistent with the findings of Glaeser et al. [1992], Henderson et al. [1995], as well as Combes [2000]. A high population density in a county leads to decrease in the growth of industry employment.[19]

[19] We considered alternative formulations, including models with the logarithm of industry (subcategory) employment as the dependent variable and the logarithm of total (green industry) employment as an explanatory variable

Table 4: Linear Regression Results for the Log of Industry Employment

| Variable | All Counties | | | | MSA Counties | | | |
|---|---|---|---|---|---|---|---|---|
| | Gr. vs. Br. (1) | | Gr. Subcats, (2) | | Gr. vs. Br. (3) | | Gr. Subcats. (4) | |
| Green industry | -5.275*** | (.293) | | | -3.398*** | (.342) | | |
| CAT1 | | | 1.689** | (.621) | | | -4.459*** | (.844) |
| CAT2 | | | -2.216*** | (.563) | | | -6.089*** | (.741) |
| CAT3 | | | -1.121** | (.532) | | | -6.358*** | (.747) |
| CAT4 | | | -1.108** | (.530) | | | -2.895*** | (.723) |
| Log (income)$_{j,t-1}$ | -.185** | (.061) | -.196** | (.079) | -.167 | (.095) | -.656*** | (.156) |
| Log (income)$_{j,t-1}\times$Green industry | .467*** | (.036) | | | .230*** | (.041) | | |
| Log (income)$_{j,t-1}\times$CAT1 | | | -.098 | (.074) | | | .595*** | (.102) |
| Log (income)$_{j,t-1}\times$CAT2 | | | .380*** | (.067) | | | .852*** | (.090) |
| Log (income)$_{j,t-1}\times$CAT3 | | | .231*** | (.064) | | | .884*** | (.091) |
| Log (income)$_{j,t-1}\times$CAT4 | | | .219*** | (.064) | | | .384*** | (.088) |
| Log (agglomeration)$_{i,j,t-1}$ | .799*** | (.016) | 1.515*** | (.017) | .975*** | (.039) | 1.336*** | (.035) |
| Log (agglomeration)$_{i,j,t-1}\times$Green industry | .088*** | (.006) | | | .211*** | (.010) | | |
| Log (agglomeration)$_{i,j,t-1}\times$CAT1 | | | -.207*** | (.013) | | | -.181*** | (.018) |
| Log (agglomeration)$_{i,j,t-1}\times$CAT2 | | | -.341*** | (.012) | | | -.304*** | (.016) |
| Log (agglomeration)$_{i,j,t-1}\times$CAT3 | | | -.295*** | (.011) | | | -.292*** | (.016) |
| Log (agglomeration)$_{i,j,t-1}\times$CAT4 | | | -.327*** | (.015) | | | -.225*** | (.019) |
| Log (agglomeration in neighbors)$_{i,j,t-1}$ | -.022* | (.011) | -.057*** | (.013) | -.200*** | (.034) | -.126*** | (.033) |
| Log (agglomeration in neighbors)$_{i,j,t-1}\times$Green industry | -.008 | (.005) | | | -.087*** | (.008) | | |
| Log (agglomeration in neighbors)$_{i,j,t-1}\times$CAT1 | | | .031** | (.011) | | | .006 | (.016) |
| Log (agglomeration in neighbors)$_{i,j,t-1}\times$CAT2 | | | .032** | (.010) | | | .004 | (.013) |
| Log (agglomeration in neighbors)$_{i,j,t-1}\times$CAT3 | | | .030** | (.009) | | | -.012 | (.013) |
| Log (agglomeration in neighbors)$_{i,j,t-1}\times$CAT4 | | | .060*** | (.012) | | | .038** | (.017) |
| Log (college funds)$_{j,t-1}$ | .000 | (.003) | -.003 | (.003) | .001 | (.002) | -.000 | (.000) |
| Log (junior college funds)$_{j,t-1}$ | -.001 | (.001) | -.002 | (.002) | -.001 | (.001) | -.002 | (.001) |
| Log (population density)$_{j,t}$ | -.151* | (.082) | -.126 | (.089) | -.215** | (.107) | -.012* | (.131) |
| Property tax rate $_{j,t}$ | .011 | (.061) | .035 | (.072) | .166 | (.132) | .214 | (.168) |
| Log (undeveloped land price)$_{j,t}$ | -.039 | (.025) | -.035 | (.030) | .021 | (.032) | -.026 | (.040) |
| Amenities employment ratio$_{j,t-1}$ | -.385** | (.173) | .074 | (.210) | -.438 | (.412)) | -.027 | (.576) |
| County effects | Yes | | Yes | | Yes | | Yes | |
| Time effects | Yes | | Yes | | Yes | | Yes | |
| Number Obs. | 14224 | | 35560 | | 4424 | | 11060 | |
| Adj. $R^2$ | .975 | | .921 | | .989 | | .952 | |

*** Denotes statistical significance at the 1% level, ** denotes statistical significance at the 5% level, and * denotes statistical significance at the 10% level.

Robust clustered standard errors are in parentheses.

The intra-green industry analysis suggests that employment growth is significantly lower in most green categories, relative to CAT5 (the omitted subcategory), conditional on the exogenous variables included in our model (the coefficient on the CAT1 dummy variable is positive in the full sample model (2) and negative in the restricted sample model (4)). Agglomeration effects are also important in explaining employment growth in the green subcategories and, although the effect is most important for CAT5, all agglomeration effects are positive and statistically significant. The agglomeration effects presented in table 3, which suggested that new firms are attracted to counties with many incumbent firms, serves to compound these effects.

### 4.3 Firm Survival

Finally we examined the survival of green and brown establishments using a proportional hazard model as proposed by Cox [1972]. Specifically, we observe the start date for each firm in our sample and infer firm exit as a disappearance of an EIN across consecutive quarters. In doing this we construct a sample of measured durations representing the time until a firm exits the market. Consider modeling the conditional probability that firm $\ell$, competing in industry $i$ and located in county $j$, exits the market at time $t$ as

$$\lambda(t|\mathbf{x}_{\ell ijt};\boldsymbol{\beta}) = \exp(\mathbf{x}'_{\ell ijt}\boldsymbol{\beta})\lambda_0(t)$$

where $\lambda(t|\mathbf{x}_{\ell ijt};\boldsymbol{\beta})$ is the conditional hazard rate and $\lambda_0(\cdot)$ is an unspecified baseline hazard function. In describing the Cox regression model, we have made a common parametric assumption that the durations follow an exponential (or, more generally a Weibull) distribution.[20]

As in Audretsch and Mahmood [1995], we restrict attention to all non-censored episodes of single birth-death transitions and allow the covariate matrix $\mathbf{X}$ to vary with time. The covariate matrix and can be divided into three groups comprising firm-level, county-level, and market-level characteristics. Firm-level covariates are the green or brown classification (or green subcategory in the intra-green industry analysis), a set of dummy variables to identify new firms and firms with past experience, the establishment's age, current average wage and relative employment compared to the county green or brown total employment. County-level variables are the agglomeration variables (within a county and considering neighboring counties), the level of university and junior college funding, population density, property tax rate, undeveloped land price, and the amenities

(the assumption being that green or brown employment differ from the trend in total employment), or models in which we tried to explain variation in the share of industry employment in the green and brown industries (within the green industry for the subcategory analysis) although the models had qualitatively similar estimation results; however, the interpretation of the coefficients differs and is more complicated given our interest in employment growth and not changes in the composition of employment which may "improve" even though employment is decreasing. These results are available from the corresponding author upon request.

[20] The approach developed by Cox [1972] is semiparametric given that the baseline hazard is "conditioned out" in estimation and need not be estimated (and thus does not require parameterization).

ratio. We also included county fixed effects as well as time fixed effects to capture market-level fluctuations.

We present the results from estimating these survival models, in the same format as our earlier estimation results, in table 5. The green-brown comparison results are, for the most part, robust to the sample definition, as can be seen by comparing columns (1) and (3). In interpreting the results, a positive (negative) coefficient indicates that the factor contributes to firms having a lower (higher) survival rate, all else equal. Thus, compared to brown industries, survival rates are lower for green industries—the failure rate of green firms is exp(0.117) = 1.124 which is 12.4% higher than than of brown firms. Similar patterns can be observed for new firms (although being a new green firm does not have any additional significant disadvantage), counties with high unemployment rates, population densities, property tax rates, and undeveloped land prices. Consistent with Dunne et al. [2005], firms with past experience survive longer and it seems experienced green firms are even more likely to remain in business. High wage rates are correlated with firm survival, and even more so for green firms. Firm agglomeration within a county is again important in explaining firm survival. The older a firm is, the more likely it is to survive and the employment ratio, which captures the size of the establishment relative to like firms in the county, indicate that larger firms survive longer.

Within the green industry, the subcategory analysis suggest mixed results which are again consistent across the different samples (which can be seen by comparing columns (2) and (4)). Relative to CAT5 firms, CAT1 firms are more likely to survive while CAT2 and CAT4 firms are less likely (conditional on all other variables). The subcategory-specific effects of being a new firm, being an experienced firm, having higher wages, and having other like firms in the industry are mixed (many of the interaction effects are small in magnitude relative to the respective total—uninteracted effect of interest or are insignificant) although the effects of these variables overall (without the interaction) are consistent with the green-brown comparison.

# 5 Characterizing Green Specialization

To provide further insight in the green economy within the State of Texas, we focused on the five subcategories and considered which locations seem suited to specialize in each subcategory given our data. To identify these comparative advantages, we calculated subcategory location quotients for each county and for each period in our data. Specifically, define

$$\mathrm{LQ}_{ijt} = \left(\frac{E_{ijt}}{E_{jt}}\right) \Big/ \left(\frac{E_{it}}{E_t}\right)$$

where we denote subcategory $i$ employment in county $j$ during period $t$ as $E_{ijt}$ and somewhat abuse subscripts in our use—dropping the $i$ subscript represents all green employment in county $j$ during

Table 5: Proportional Hazard Model Estimates

| Variable | All Counties | | | | MSA Counties | | | |
|---|---|---|---|---|---|---|---|---|
| | Gr. vs. Br. (1) | | Gr. Subcats. (2) | | Gr. vs. Br. (3) | | Gr. Subcats. (4) | |
| Green industry | .117** | (.037) | | | .153** | (.041) | | |
| CAT1 | | | -.354** | (.116) | | | -.383** | (.132) |
| CAT2 | | | .320** | (.079) | | | .312** | (.085) |
| CAT3 | | | -.032 | (.076) | | | -.019 | (.083) |
| CAT4 | | | .302** | (.100) | | | .230* | (.123) |
| New firm | .558** | (.006) | .466** | (.015) | .550** | (.006) | .471** | (.016) |
| New firm× Green industry | .003 | (.007) | | | .012 | (.007) | | |
| New firm×CAT1 | | | -.009 | (.019) | | | -.030 | (.021) |
| New firm×CAT2 | | | .060** | (.013) | | | .046** | (.014) |
| New firm×CAT3 | | | .071** | (.013) | | | .065** | (.013) |
| New firm×CAT4 | | | -.088** | (.018) | | | -.052** | (.021) |
| Firms with past experience | -.102** | (.003) | -.104** | (.010) | -.102** | (.003) | -.104** | (.011) |
| Firms with past experience×Green industry | -.049** | (.007) | | | -.048** | (.007) | | |
| Firms with past experience×CAT1 | | | -.012 | (.019) | | | -.009 | (.021) |
| Firms with past experience×CAT2 | | | -.063** | (.013) | | | -.059** | (.014) |
| Firms with past experience×CAT3 | | | -.025** | (.013) | | | -.021 | (.013) |
| Firms with past experience×CAT4 | | | .028** | (.017) | | | .026 | (.020) |
| Log (wage)$_{\ell,t}$ | -.032** | (.002) | -.035** | (.007) | -.032** | (.002) | -.038** | (.007) |
| Log (wage)$_{\ell,t}$× Green industry | -.014** | (.004) | | | -.018** | (.005) | | |
| Log (wage)$_{\ell,t}$×CAT1 | | | .047** | (.013) | | | .049** | (.015) |
| Log (wage)$_{\ell,t}$×CAT2 | | | -.034** | (.009) | | | -.032** | (.009) |
| Log (wage)$_{\ell,t}$×CAT3 | | | .004 | (.009) | | | .001 | (.009) |
| Log (wage)$_{\ell,t}$×CAT4 | | | -.040** | (.011) | | | -.035** | (.014) |
| Log (agglomeration)$_{i,j,t-1}$ | -.054** | (.003) | -.067** | (.009) | -.064** | (..003) | -.086** | (.010) |
| Log (agglomeration)$_{i,j,t-1}$×Green industry | -.003 | (.006) | | | -.008 | (.006) | | |
| Log (agglomeration)$_{i,j,t-1}$×CAT1 | | | -.056** | (.017) | | | -.056** | (.020) |
| Log (agglomeration)$_{i,j,t-1}$×CAT2 | | | .005 | (.011) | | | .007 | (.012) |
| Log (agglomeration)$_{i,j,t-1}$×CAT3 | | | .022** | (.010) | | | .022* | (.011) |
| Log (agglomeration)$_{i,j,t-1}$×CAT4 | | | .016 | (.015) | | | .028 | (.018) |
| Log (agglomeration in neighbors)$_{i,j,t-1}$ | .043** | (.003) | .049** | (.009) | -.140** | (.030) | .063** | (.010) |
| Log (agglomeration in neighbors)$_{i,j,t-1}$×Green industry | .006 | (.005) | | | -.010** | (.006) | | |
| Log (agglomeration in neighbors)$_{i,j,t-1}$×CAT1 | | | .028* | (.016) | | | .029 | (.019) |
| Log (agglomeration in neighbors)$_{i,j,t-1}$×CAT2 | | | .002 | (.010) | | | .002 | (.011) |
| Log (agglomeration in neighbors)$_{i,j,t-1}$×CAT3 | | | -.013 | (.010) | | | -.010 | (.011) |
| Log (agglomeration in neighbors)$_{i,j,t-1}$×CAT4 | | | -.002 | (.014) | | | -.009 | (.017) |
| Log (age)$_{\ell,t}$ | -.420** | (.002) | -.429** | (.005) | -.424** | (.002) | -.434** | (.005) |
| Employment ratio$_{\ell,t}$ | -.107** | (.005) | -.093** | (.011) | -.140** | (.006) | -.119** | (.012) |
| Log (college funds)$_{j,t-1}$ | -.008** | (.000) | -.007** | (.001) | -.010** | (.000) | -.009** | (.001) |
| Log (junior college funds)$_{j,t-1}$ | -.001* | (.000) | .001 | (.001) | -.000 | (.000) | .001 | (.001) |
| Unemployment rate$_{j,t-1}$ | .005** | (.001) | .008** | (.002) | .006** | (.000) | .009** | (.002) |
| Log (population density)$_{j,t-1}$ | .004** | (.002) | .006* | (.004) | .020** | (.003) | .019** | (.005) |
| Property tax rate $_{j,t-1}$ | .060** | (.014) | .056** | (.027) | .009 | (.017) | -.087** | (.034) |
| Log (undeveloped land price)$_{j,t-1}$ | .190** | (.003) | .179** | (.005) | .229** | (.003) | .224** | (.007) |
| Number Obs. | 599818 | | 143610 | | 521173 | | 122051 | |
| Wald $\chi^2$ | 188405.44 | | 43416.69 | | 171004.15 | | 39213.27 | |

** Denotes statistical significance at the 5% level and * denotes statistical significance at the 10% level.

Robust clustered standard errors are in parentheses.

period $t$ ($E_{jt}$), while dropping the $j$ subscript represents total subcategory $i$ employment throughout Texas during period $t$ ($E_{it}$), and dropping both the $i$ and $j$ subscript denotes all green employment in the state during period $t$. Location quotients are often used to compare the economy of one region to that of some benchmark economy (which typically encompasses the regional economy). In our case, we compare the share of green employment in each subcategory to that share of the state's green employment working in that subcategory. To summarize these measures in a concise way, we compute a simple average over the data periods as

$$\mathrm{LQ}_{ij} = \frac{1}{T}\sum_{t=1}^{T}\mathrm{LQ}_{ijt}$$

and interpret this as county $j$'s location quotient in subcategory $i$. A value of one suggests that the employment share is exactly in line with that of the state employment mix (on average). Location quotients are often used to characterize whether a region is an exporter (importer) of the product produced in a given industry based on whether its value is sufficiently greater (less) than one. It is in this sense that we appeal to location quotients to help identify which counties might have a comparative advantage in certain green tasks.

In figure 4, we depict a map of the State of Texas in which we have shaded each county one of five colors. The colors correspond to the five green subcategories and the color choice was made by selecting the maximum of each county's $\mathrm{LQ}_{ij}$ values; i.e., $\max\left\{\mathrm{LQ}_{1j}, \mathrm{LQ}_{2j}, \mathrm{LQ}_{3j}, \mathrm{LQ}_{4j}, \mathrm{LQ}_{5j}\right\}$. Clear patterns arise that are consistent with what a state planner might suggest: counties with large wind farms (such as Nolan and Upton counties—denoted in the map by the Abilene and Odessa MSAs, respectively) specialize in CAT1; population-dense counties (such as those around major metropolitan areas—in particular, the Dallas–Fort Worth area) often specialize in CAT2 which primarily concerns LEED-certified or LEED-eligible buildings or construction; the activities in CAT3 seem more diverse and hence there is more variation in the driving factors leading to specialization within this green subcategory; cattle production accounts for nearly half of the state's cash receipts for agricultural commodities and the state's three largest cattle-producing counties (Deaf Smith and Castro counties—denoted in the map by the areas around the Amarillo MSA) as well as most rural counties (most of West Texas) all specialize in CAT4 which encompasses NAICS codes concerning organic meat and produce; counties containing colleges and universities (such as the University of Texas in Travis County—denoted by the Austin MSA, Texas A&M University in Brazos County—denoted by the College Station MSA, and Texas Tech University in Lubbock County—denoted by the Lubbock MSA) often specialize in CAT5.

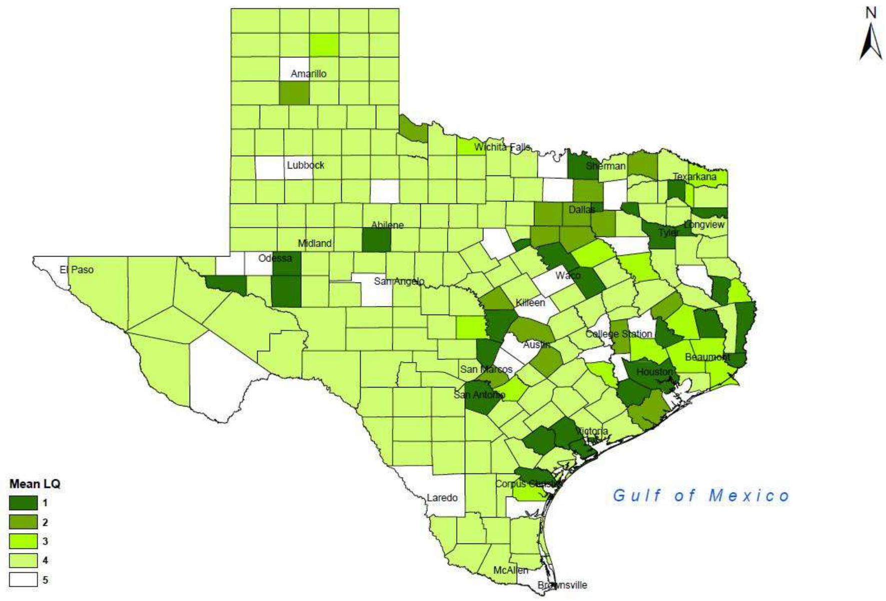



Figure 4: Maximum of Subcategory Location Quotient Averaged over Periods

# 6 Conclusion

We took as given the BLS definition of the green industries and compared features of these industries to that of all other (brown) industries over a data period that took place before the definition was developed and within the State of Texas. In general, we find little evidence that the factors that encourage entry, growth, or exit are inherently different for green and brown industries. Although there are some specific effects that can be highlighted (for example, income effects are important for employment growth in green, but not brown, industries), agglomeration within a county seems to be the primary factor in attracting, growing, and retaining the green industries. Within the green industry, our subcategory analysis highlights more distinctions across the types of green activities. A caveat of taking the BLS definition as given is that the green versus brown industries may involve too much aggregation, although there is little we can do to address this without challenging the stance taken on specific NAICS codes that comprises the BLS definition. The definition, however, provides as much detail as can be hoped for under NAICS (six digits are the finest level of detail that NAICS allows).

Our investigation of the non-policy factors that affect the green industry cannot speak to the efficacy of various policies designed to attract such firms. For example, Palmer and Burtraw [2005] compared policies aimed at increasing the contribution of renewables to the total U.S. electricity supply. It would be interesting to investigate if these policies (such as production tax credits, renewable portfolio standards, cap-and-trade policies, etc.) are successful in attracting firm investment and growing green industries relative to the non-policy factors we've investigated. In some sense, our results offer two potential interpretations: first, attracting green investment is no different than attracting non-green investment; alternatively, attracting just a little bit of green investment can pay large dividends in that agglomeration effects can obtain in proliferating the green industry. Our initial characterization of the green industry was restricted to an area in which policy differentials were nonexistent which prevents us from distinguishing between these two interpretations. However, one could leverage variation in policies across states (in either a subnational, or international evaluation, given most environmental policies are set at the federal level) to consider the costs and benefits of policies aimed at encouraging sustainable development. Such policy-specific research would provide a valuable contribution in complementing our general analysis. Evaluating such policies could consider the effects on individual industries, one of the green subcateogries, or the green industry as a whole and can consider analysis across regions, states, or countries. Given our findings suggest that green and brown industries do not respond in drastically different ways to non-policy factors, this may suggest a role for policymakers hoping to stimulate green investment and growth.

## Acknowledgements

We thank Tim Dunne as well as participants at the 2010 Western Economic Association International conference and the 2011 Midwest Economic Association conference for helpful comments. We also thank seminar participants at Oberlin College and the environmental economics workshop at Texas Tech University. We acknowledge the U.S. Department of Energy (award number DE-FG36-06GO86092) for financial support and the Texas Workforce Commission for providing the data.

# References Cited

Laura Abramovsky, Rupert Harrison, and Helen Simpson. University Research and the Location of Business R&D. *Economic Journal*, 117(519):C114–C141, 03 2007.

David B. Audretsch and Talat Mahmood. New Firm Survival: New Results Using a Hazard Function. *The Review of Economics and Statistics*, 77(1):97–103, 1995.

Anthony J. Barbera and Virginia D. McConnell. The Impact of Environmental Regulations on Industry Productivity: Direct and Indirect Effects. *Journal of Environmental Economics and Management*, 18(1):50–65, 1990.

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

Smita B. Brunnermeier and Mark A. Cohen. Determinants of Environmental Innovation in US Manufacturing Industries. *Journal of Environmental Economics and Management*, 45(2):278–293, 2003.

A. Colin Cameron and Pravin K. Trivedi. *Microeconometrics: Methods and Applications*. Cambridge Books. Cambridge University Press, 2005.

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

Pierre-Philippe Combes. Economic Structure and Local Growth: France, 1984-1993. *Journal of Urban Economics*, 47(3):329–355, 2000.

David R. Cox. Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2):187–220, 1972.

Dakshina G. De Silva and Robert McComb. Research Universities and Regional High-Tech Firm Start-Up and Exit. *Economic Inquiry*, forthcoming. doi: 10.1111/j.1465-7295.2010.00353.x.

Michael P. Devereux, Rachel Griffith, and Helen Simpson. Firm Location Decisions, Regional Grants and Agglomeration Externalities. *Journal of Public Economics*, 91(3-4):413–435, 2007.

Timothy Dunne, Mark J. Roberts, and Larry Samuelson. Patterns of Firm Entry and Exit in U.S. Manufacturing Industries. *RAND Journal of Economics*, 19(4):495–515, 1988.

Timothy Dunne, Mark J. Roberts, and Larry Samuelson. The Growth and Failure of U.S. Manufacturing Plants. *The Quarterly Journal of Economics*, 104(4):671–698, 1989.

Timothy Dunne, Shawn D. Klimek, and Mark J. Roberts. Exit from Regional Manufacturing Markets: The Role of Entrant Experience. *International Journal of Industrial Organization*, 23 (5-6):399–421, 2005.

Piet Eichholtz, Nils Kok, and John M. Quigley. Doing Well by Doing Good? Green Office Buildings. *American Economic Review*, 100(5):2492–2509, 2010.

Edward L. Glaeser, Hedi D. Kallal, Jose A. Scheinkman, and Andrei Shleifer. Growth in Cities. *Journal of Political Economy*, 100(6):1126–1152, 1992.

Christian Gourieroux, Alain Monfort, and Alain Trognon. Pseudo Maximum Likelihood Methods: Applications to Poisson Models. *Econometrica*, 52(3):701–20, 1984.

Paulo Guimarães. The Fixed Effects Negative Binomial Model Revisited. *Economics Letters*, 99 (1):63–66, 2008.

J. Vernon Henderson. Efficiency of Resource Usage and City Size. *Journal of Urban Economics*, 19(1):47–70, 1986.

Vernon Henderson, Ari Kuncoro, and Matt Turner. Industrial Development in Cities. *Journal of Political Economy*, 103(5):1067–1090, 1995.

Arik Levinson. Environmental Regulations and Manufacturers' Location Choices: Evidence from the Census of Manufactures. *Journal of Public Economics*, 62(1-2):5–29, 1996.

John A. List. US County-Level Determinants of Inbound FDI: Evidence from a Two-Step Modified Count Data Model. *International Journal of Industrial Organization*, 19(6):953–973, 2001.

John A. List and Catherine Y. Co. The Effects of Environmental Regulations on Foreign Direct Investment. *Journal of Environmental Economics and Management*, 40(1):1–20, 2000.

Karen Palmer and Dallas Burtraw. Cost-Effectiveness of Renewable Electricity Policies. *Energy Economics*, 27(6):873–894, 2005.

Stuart S. Rosenthal and William C. Strange. Geography, Industrial Organization, and Agglomeration. *The Review of Economics and Statistics*, 85(2):377–393, 2003.

Tim Simcoe. XTPQML: Stata Module to Estimate Fixed-Effects Poisson (Quasi-ML) Regression with Robust Standard Errors. Statistical Software Components, Boston College Department of Economics, 2007.

Udo Staber. Spatial Proximity and Firm Survival in a Declining Industrial District: The Case of Knitwear Firms in Baden-Württemberg. *Regional Studies*, 35(4):329–341, 2001.

Douglas Woodward, Octávio Figueiredo, and Paulo Guimar aes. Beyond the Silicon Valley: University R&D and High-Technology Location. *Journal of Urban Economics*, 60(1):15–32, 2006.

Jeffrey M. Wooldridge. Distribution-Free Estimation of Some Nonlinear Panel Data Models. *Journal of Econometrics*, 90(1):77–97, 1999.

# A Data Appendix

As we discussed in describing the data we used in our research, one difficulty we faced concerned the mapping between the 2002 NAICS codes which characterize the QCEW data and the 2007 NAICS codes which were used by the BLS to define the green categories. To identify the green categories in the 2002 NAICS classification we used the concordances provided by the U.S. Census Bureau—see footnote 10. We were forced to drop seven 2002 NAICS industries which appear to be extinct—none of the industries had any new firms enter over our sample period and are not listed in the concordances.[21] In situations where one 2002 NAICS code corresponded with multiple 2007 NAICS codes, we specified the industry as green if and only if all 2007 NAICS codes were defined to be green. In this sense our definition of the green industry is slightly more conservative than the definition intended by the BLS. Specifically, because of this reason we were forced to drop eleven 2002 NAICS codes due to conflicts in the way the corresponding 2007 NAICS codes were classified—at least one 2007 NAICS code was defined as green while at least one other code was considered non-green.[22] A related issue concerned defining which 2002 NAICS sectors comprised the green category variables. For instances in which a 2002 NAICS code mapped to multiple 2007 NAICS codes, we assigned the set of all the 2007 NAICS subcategories to the 2002 NAICS code. For example, 2002 NAICS code 235710 is associated with 2007 NAICS codes 238991, 238992, 238111, and 238112, which are all considered green industries under the BLS definition. Industries 238111 and 238112 are considered part of CAT2 and all four industries are considered part of CAT3 and CAT4 under the BLS definition. Thus, we defined the 2002 NAICS code 235710 which encompasses these four 2007 NAICS sectors to be part of CAT2, CAT3, and CAT4.

In table A.1, we describe the covariates used in the models we considered.

[21] Specifically, we dropped the 2002 NAICS codes 239891, 502362, 505221, 505411, 505413, 508141, and 513330.

[22] For this reason, we had to drop 2002 NAICS codes 234120, 234910, 234930, 235210, 315211, 315212, 326291, 326299, 339111, 421930, and 514199.

Table A.1: Conditioning Variables

| Variable | Description |
|---|---|
| BLS Green | An indicator variable corresponding to whether the observation relates to the green (= 1) or brown (= 0) industry. |
| CAT1 | An indicator variable corresponding to whether the observation was identified as relating to "renewable energy" (= 1) or not (= 0). |
| CAT2 | An indicator variable corresponding to whether the observation was identified as relating to "energy efficiency" (= 1) or not (= 0). |
| CAT3 | An indicator variable corresponding to whether the observation was identified as relating to "pollution abatement and/or recycling" (= 1) or not (= 0). |
| CAT4 | An indicator variable corresponding to whether the observation was identified as relating to "natural resource conservation" (= 1) or not (= 0). |
| CAT5 | An indicator variable corresponding to whether the observation was identified as relating to "environmental compliance, training, and awareness" (= 1) or not (= 0). |
| Income | Average wages per quarter paid in the county for all establishments as reported in QCEW data. |
| Agglomeration | The number of like firms already present in the county of a certain type (i.e., green, brown, or belonging to a specific category.) |
| Agglomeration in neighbors | Analogous to agglomeration, but considers like firms in all contingent counties for a given (base) county. |
| College funds | University and research center R&D expenditures reported by the NSF. The annual NSF data actually span two calendar years. To convert these annual R&D expenditures into quarterly data, we used a fourth of a given year's total for Q1–Q3 (each), and a fourth of the given year's total for Q4 of the previous calendar year (as the federal fiscal year begins in October). Although the NSF provides research funding by institution and is identified by granting agency (i.e., DoE, EPA, DoD, etc.) we aggregated total federal awards by geographically distinct institutions (i.e., system campuses are scored separately) to compute a measure of R&D at the county level. |
| Junior college funds | The same as college funds, but R&D expenditures by junior colleges within a county for a given quarter. |
| Unemployment rate | The county-specific, seasonally unadjusted, end-of-quarter unemployment rate. |
| Population density | County-estimated annual population density from U.S. Census. |
| Property tax rate | Annual county property tax rate. |
| Undeveloped land price | The yearly, median undeveloped land price in each of 33 land market regions in Texas for the counties comprising the region, as reported by the Texas A&M Real Estate Center. |
| Amenities employment ratio | The share of county employment in NAICS 71, (arts, entertainment, and recreation), NAICS 721110 (hotels and motels), 722110 (full service restaurants), and 722410 (drinking places, alcoholic beverages) as reported in the QCEW data. The NAICS 72 activities also reflect the scope of the locality's amenities for business travelers as well as informal business and social interaction. |
| Age | The number of months that have passed since the firm's start date. |
| New firm | Firms that are three years old or younger (have an age of no more than 36 months). |
| Firms with past experience | Firms that have a prior existing EIN and were later reassigned; for example, if the establishment changed hands or if a partnership was broken up. We do not observe the reason for the change. |
| Wage | Establishment-level quarterly average wage which was calculated by dividing the quarterly wage bill for the establishment by its average number of employees for that quarter. |
| Employment ratio | The firm's current employment as a share of total industry (or subcategory) employment within a county for a given quarter. |