The Troubled Asset Relief Program's Effect on Firm Entry over the Business Cycle

Kevin D. Duncan

March 5, 2019

Abstract

This paper studies the impacts of banks receiving investment from the Treasury Department's Capital Purchase Program (CPP) on surrounding county business dynamics during the 2008 Financial Crisis. I hypothesize prospective firms had an easier time gaining loans in counties that received CPP funds relative to similar counties that did not, resulting in higher firm entry in treated counties. I estimate the average treatment for the treated of a county receiving CPP funds on the log number of new establishments in the 4 years following CPP injections. My estimation strategy combines several previously disjointed matching estimator techniques, including dynamic matching techniques ([Lechner, 2008]) while relaxing the traditional Stable Unit Treatment Value Assumption (SUTVA) to account for interference effects ([Liu et al., , Forastiere et al., 2018]). I further utilize this structure to discuss synthetic control estimators ([Abadie et al., 2010]) without the SUTVA assumption. This structure enables estimation of both the main effect of a county receiving treatment, as well as spillover effect of neighbors. Preliminary results find no discernible difference in county business dynamics over this time horizon between treated an comparable untreated counties.

1 Introduction

This paper examines the role of Treasury Department's Capital Purchase Program (CPP) on firm entry during the 2008 financial crisis. The Capital Purchase Program provided $205 billion dollars to banks in order to stabilize banks and stimulate loan supply as part of the broader Troubled Asset Relief Program (TARP). I utilize matching and synthetic control estimators extended to account for spatial correlation between counties and the dynamic assignment of treatment of CPP funds to show that the Average Treated on the Treated is indistinguishable from zero. This null result is notable for a few reasons. Firm entry over the business cycle is a main contributor to TFP growth, despite being usually counter-cyclical ([Lee and Mukoyama, 2015], [Clementi and Palazzo, 2016], [Gourio et al.,]). Higher entry rates tend to lead to lower unemployment and stronger economic growth out of economic depressions.

Previous studies have focused on bank level response to receiving CPP funds, however considerable heterogeneity exists between type of banks and lending behavior. Trying to provide funds directly to banks is similar to the pass through of monetary policy changes to credit markets. However, banks show heterogeneous responses to monetary policy shocks ([Blau et al., 2013]). Secondly, most small businesses do not have access to equity markets, and rely on local or regional banks for credit. Relationship lending has been recently established as a major way in which banks recover underlying firm specific behavior ([Berger and Udell, 2002]). However, this relationship lending is often only carried out at the request of loan offers, such that regional and local bank structure might continue to create heterogeneous response across banks in response to receiving CPP funds. Comparably this paper focus on how the injection of CPP funds to local communities affected county level firm entry. I view the transition mechanism as banks received CPP funds, and if this represented a relaxation of local credit constraints this should have led to higher relative firm entry among participating banks. This allows a secondary means of carrying out program evaluation that policy makers might judge program efficiency on.

My main departure from the existing literature is that I focus on firm entry over behavior of banks that received CPP funds. Many of these studies have come to inconclusive if not contradictory results. Among studies that have focused on if the CPP led to increased bank lending, many come to inconclusive results. [Contessi and Francis, 2011] examine how the CPP impacted total loans, real estate loans, commercial and industrial loans, and consumer loans and find no difference in their business and consumer loaning behavior. [Cole, 2012] uses a panel fixed effects model to show that banks receiving CPP funds failed to increase their small business loans. [Lei, 2013] instrument the probability of a bank receiving TARP funds by their political and regulatory connections and estimate banks that received TARP funds increased their loan supply by 6.36% of total assets. [Bassett, 2016] use the same instruments as [Lei, 2013] to show banks did not change the growth rate of loans.

Analysis of the CPP benefits from several stylized facts; the CPP had statutory requirements where the Treasury could only purchase stock valued between 1%-3% of a banks troubled assets, up to $25 billion. The CPP provided money to 708 banks in 462 counties, and among counties that received money, only a few banks received CPP funds. Most importantly, the funds were dispersed quickly, with all the initial money poring into banks between November 2008 and December 2009. Many of the counties that had participating banks received injections only during specific periods of the capital outlay. These features allow me to view as counties as either 'treated' or 'non-treated,' over only a few time periods, versus a more complex problem with multiple treatment levels as a continuous treatment process.

I focus on generating estimates for the Average Treatment for the Treated (ATT) represented by the change in the log number of establishments in continental US counties that contained eligible banks. I focus on calculating the ATT due to the heterogeneity in location responses to the financial crisis. States such as New Jersey, Florida, and Nevada were hit the hardest, while Midwest states saw very little decline in local housing prices and employment. Estimates of the ATE would estimate counterfactuals for states whose local credit market conditions did not deteriorate over the 2008 financial crisis, and therefore naturally draw estimates towards zero. Traditional estimators of these effects are ideal for independent observations. Estimating the effect on counties requires extending both matching and synthetic control estimators to be able to account for both dynamic treatment assignment and spillovers due to neighbor treatment status.

These estimators build on combining several established methods and showing under appropriate conditions they still function as intended. Dynamic treatment effects has been covered both by inverse propensity score weighting in Marginal Structural Models ([Robins et al., 2000, Imai and Ratkovic, 2015]), and matching estimators ([Lechner, 2005, Lechner, 2008]). Estimators for interference effects [Sobel, 2006, Hudgens and Halloran, 2008, Tchetgen Tchetgen and VanderWeele, 2012, Liu et al., , Forastiere et al., 2018]. Similar extensions to synthetic control estimators with interference effects have recent contributions ([Cao and Dowd, 2019]). The methods provide data-driven methods to estimate counterfactuals when natural counterfactuals are not clear. Each of the methods face their own trade-offs. Dynamic matching or IPW requires estimation of a multi-stage propensity score based around multiple treatment and neighbor characteristics. The resulting estimates are the mean effect. Synthetic control estimates allow estimation of individual specific effects, at the cost of assuming a linear response function.

Identifying this impact faces possible endogeneity in choices of the primary regulators over which banks received CPP funds due to unobserved loan demand characteristics in receiving counties. Regulators might have placed preference over accepting banks that met a variety of objectives that are correlated with potential outcomes. For example, regulators may have approved banks for TARP funds to either areas with weak or high loan demand. Areas with weak loan demand were correlated with banks most likely to be at risk, while areas with high, but unserved, loan demand might be desirable for potential higher firm entry. I use the same instruments as [Lei, 2013] to break up this endogeneity by using measures of local political connectedness. [Blau et al., 2013] examine the chance of US banks receiving TARP money based on a set of measures of political connectedness, and find that not only do they receive it more often, but receive more money on less costly conditions. The resulting exclusion restrictions imposed are that regulatory and political connections of counties impacted bank's chance of receiving TARP funds, and that these connections were independent of local loan demand.

Section 2 discusses the Capital Purchase Program, Section 3 discusses my identification strategy, section 4 provides information on my data sources and preliminary analysis, Section 5 generates estimation results, and Section 6 concludes. My results indicate that the CPP did not impact subsequent firm entry in treated counties. Rather than take a model average approach, holistically I employ a variety of estimators which should approximate the true ATT. I recover many estimates close to zero, with different signs but equal magnitude across a variety of time periods as equivalent levels of statistical significance.

2 The Capital Purchase Program

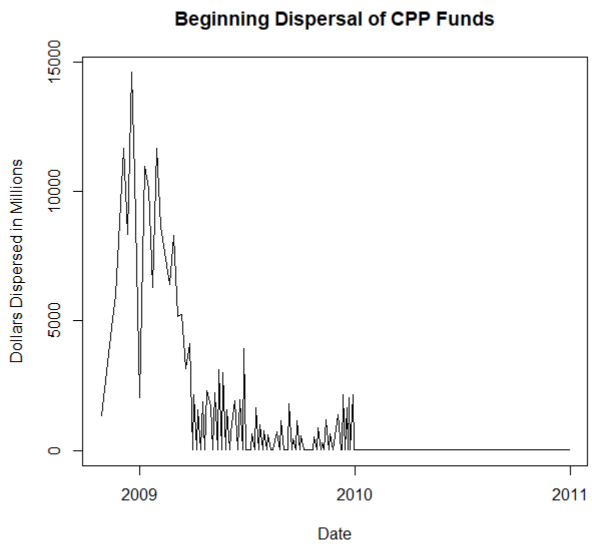
The Capital Purchase Program aimed to provide an extra layer of capital for stability or lending by buying up non-voting senior preferred shares on standardized terms across all participating banks. The maximum amount of capital eligible for purchase by the Treasury under the CPP was the lesser of (i) an amount equal to 3 percent of the Total Risk-Weighted Assets of the applicant or (ii) $25 billion. The minimum amount eligible for purchase under the CPP was the amount equal to 1 percent of the Total Risk-Weighted Assets of the applicant. All measurements were based on the information contained in the latest quarterly supervisory report filed by the applicant updated to reflect events materially affecting the financial condition of the applicant occurring since the filing of such report.

These purchases were not cost free to participation banks, requiring a 5% dividend for the first 5 years, and 9% afterwards. Participating banks would also gain 10 year warrants on future Treasury purchases of common stock that could buy up to 15% of the initial CPP investment. Despite this, research has indicated that these purchases were still generally preferential for the banks. The Congressional Oversight Panel estimated that across all TARP programs, the Treasury gave out $254 billion in 2008, for which it received assets worth approximately $176 billion, a deficit of $78 billion. Equivalently, [Veronesi and Zingales, 2010] estimate during the first 10 transactions of the CPP, the Treasury overpaid between $36-13 billion for financial claims.

Individual banks applied for CPP funds through their federal regulator- the Federal Reserve, FDIC, Office of the Comptroller of the Currency, or the Office of Thrift Supervision. Banks include the number of preferred shares they want, amount of authorized but unissued preferred stock available, authorized but unissued common stock, amount of total risk-weighted assets as reported on recent FR-Y9, call report, or TFR, why the bank might not be fully compliant before Nov 14, 2008, a time line for becoming compliant, and the type of bank. If the applicant was a bank holding company, the application was submitted to both the applicant's holding company supervisor and the supervisor of the largest insured depository institution controlled by the applicant.

The federal regulatory agency then got to choose which banks were to receive money, and sent their preferred set of applicants to the Treasury Department for final clearance. Of importance is that applications that were rejected or withdrawn were not announced or publicly disclosed.[1] The application period lasted between October 3rd, 2008 to November 14th, 2008 for publicly held companies, December 8th for Privately held companies, and February 13th, 2008 for S Corporations. There was only the single application period. Despite this, there is considerable differences in the timing of TARP injections.

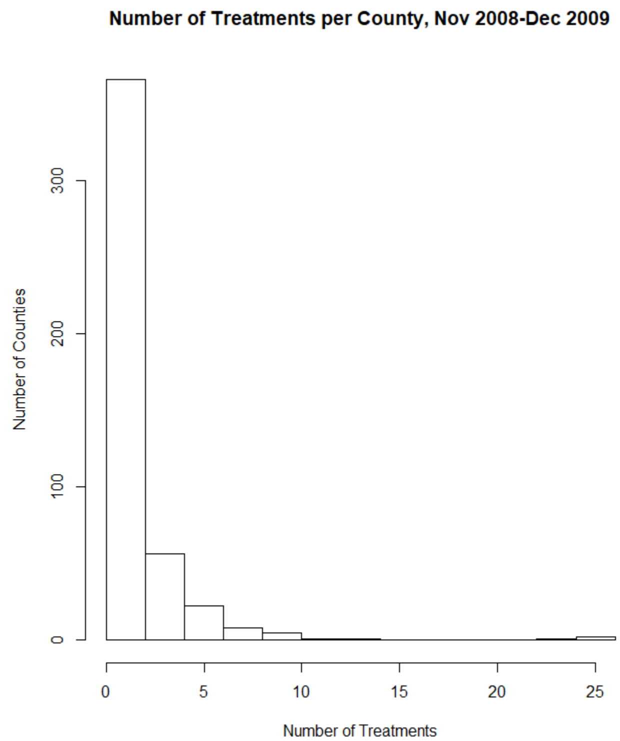
Payments began almost immediately to the most distressed banks- usually the largest. All payments to participating banks were made before January 1st, 2010. However, there are clear spikes in lending. Between November 14th and Dec 31st 2008, there is a clear peak in CPP payments to banks. There was then a second peak at the beginning of 2009. On May 20th, 2009, Timothy Geithner announced that for banks with assets less than $500 million would have a second window to apply for CPP funds for the following 6 months.[2] This generated a significantly lower, but still residual payment stream for the remainder of 2009.



Many counties received multiple injections of funds. Between 2008 and the end of 2010, the average county that received treatment, received 2.06 injections. Over the course of the program's history, among counties that received CPP funds, the average number of treatments per county was 5.147.

[1] [Duchin and Sosyura, 2014] show that of roughly 600 public firms, 416 firms (79.8%) applied, 329 (79.1%) were accepted, and that 278 (84.5%) accepted the funds but 51 (15.5%) declined. The population of private bank's behavior is unknown, and I act as if every bank that was offered funds received them.

[2] https://www.treasury.gov/press-center/press-releases/Pages/tg139.aspx



The result of this program structure favors viewing the program as the government investing largely into counties where credit market conditions had deteriorated. I view a county being treated as a binary event in each time period as long as at least one bank in a county received CPP funds. Figure 1 shows what counties in 2008 received CPP funds, and the initial dispersion is heavily biased to large urban areas on the East and West coast. Many of the counties that received grants in this time period also contained the headquarters of the largest bank or bank holding companies in the US. Equivalently Figure 2 shows which counties received CPP funds in 2009. CPP funds are far more disperse throughout 2009, however these locations are still heavily tied to major metropolitan areas in the US.

In the next section I describe what data and variables I use to estimate the average increase in firm entry rate among treated counties. The main issues that my identification strategy has to address includes both the dynamic assignment of treatment, spill overs in regional credit markets that might extend beyond the treated county, and possible spatial correlation in the errors between counties.

3 Identification Strategy

I want to estimate the average increase in firm entry rates due to at least one bank in a given county receiving CPP funds, among counties that received CPP funds. I am interested in estimating this average increase in firm entry rates for the three or four years following a county receiving CPP funds. Traditional estimators might not capture the full effect due to the presence of policy spill overs from neighboring counties treatment status, and the dynamic treatment assignment. Thus I want to first create a framework to think about spill

Figure 1: Map of all Counties treated in 2008

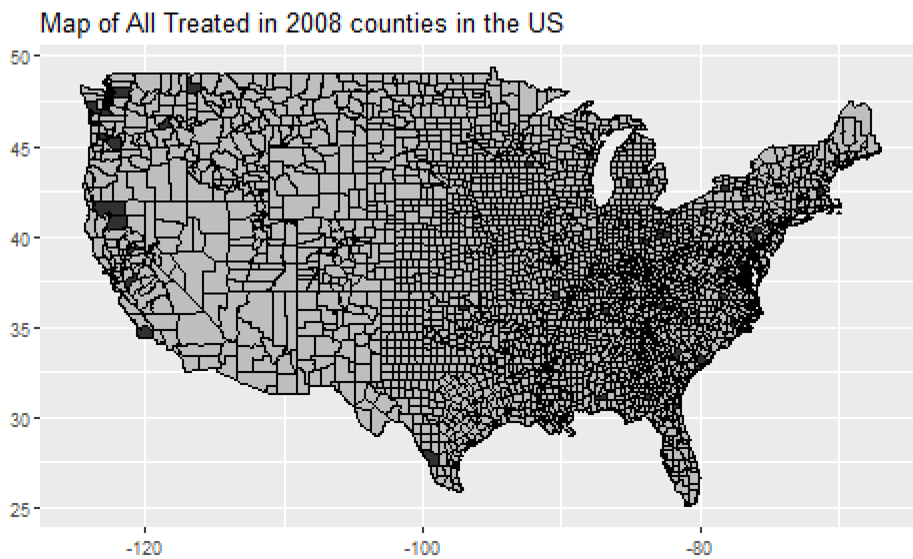
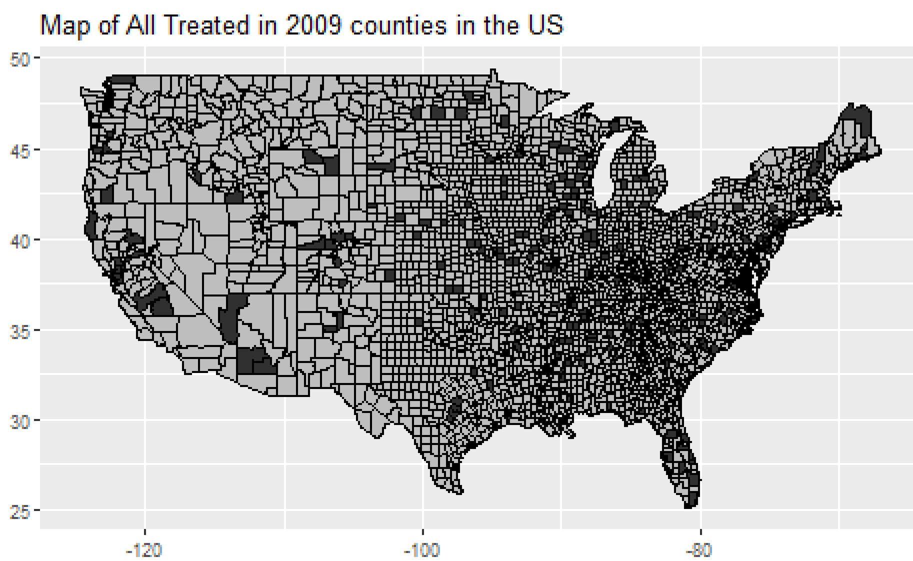


Figure 2: Map of all Counties treated in 2009



overs between treated units, while incorporating a dynamic treatment policy.

Let there exist a lattice $D \subset \mathbb{R}^2$ is infinite countable. All elements in D are located at distances of at least $d_0 > 0$ from each other, i.e. $\forall i, j \in D; \rho(i,j) \geq d_0 > 0$. For any n, define $D_n \subset D$. This generalizes the usual sample to be over an unknown, infinite lattice, where as n increases, the number of observations on the lattice grows. These points define county centroids in practice, and without loss of generality, let d_0 be the smallest distance between any two county center's in the US. The assumption of a minimum distance ensures unbounded expansion of sample regions, and rules out in-fill asymptotics.

There are four time periods, which I index by $t = 0, 1, 2, 3$. In period 0 each county receives no treatment, in period's 1 and 2, some counties receive treatment, and in period 3 no additional treatments are assigned. The requirements of 4 time periods can be relaxed such that there are infinite numbers of pre-treatment time periods, as well as a large number of post-treatment time periods. However, treatment status in the pre and post treatment periods is the same across all observations. As a result in period 1 a county can be observed in exactly one of two treatments $(0, 1)$. In period 2 she participate in one of *four* treatment sequences $((0,0),(1,0),(0,1),(1,1))$. The remaining history is identical across all counties, so can be dropped from their potential outcomes. Define $T_0((\underline{s_t})$ as the *last* time a county received treatment under treatment history $(\underline{s_t})$.

Define the vector of random variables $S = (S_0, S_1, S_2, S_3)$ to be the treatment received by a member of the population. A particular realization of S_t is denoted by $s_t \in \{0, 1\}$. Then define the history of treatment up to period t by a bar below a variable, e.g. $\underline{s_2} = (s_1, s_2)$. This allows tracking of which history a county is following, and appropriate indexing for estimating Average Treatment Effects. This is still not enough in to define the potential outcomes of each county. Since there are plausible spill overs from neighbors, I need to link neighbor treatment status to own potential outcomes. To do this I next define how each node relates to each other. This is done by describing a simple graph, or network, between points.

Assumption 3.1. *For every n, there exists a binary relation on D_n denoted W_n which is an $n \times n$ matrix showing county adjacency. The pair (D_n, W_n) is called a non-directional graph.*

Since the graph is assumed to be non-directional, if node a is connected to node b, then b is connected to a. This further implies that W_n is symmetric. I only require the existence of at least one known W_n, but does not rule out multiple W_n. For example, a common adjacency matrix is defined by borders that share contiguous borders. In this context even if two points $(a, b) \in D_n$ are adjacent, and share the same distance from a third county centroid $c \in D_n$, there is no guarantee that (a, b) share borders with c.

To motivate a simple example of different graphs, Figure 3 plots the five counties of Rhode Island and each county's centroids. There are now a few methods that I can use to generate adjacency matrices. Defining adjacency to be based on border contiguity, I graph the resulting border congruency matrix in Figure 4, and the matrix representation of this graph is shown in Table 1.

If instead I wish to use a county-centroid metric of adjacency, I first plot the distance between all counties in Table 2. The matrix represents the distance between counties in miles. Now, I pick a distance

Figure 3: Rhode Island County Borders

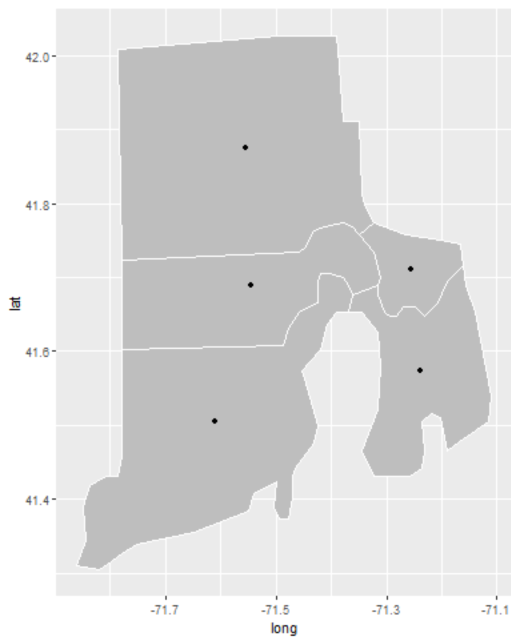


Figure 4: Border Adjacency Network

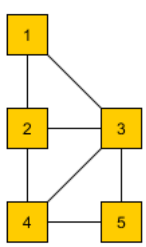


Table 1: Border Adjacency Matrix

	1	2	3	4	5
1	0	1	1	0	0
2	1	0	1	1	0
3	1	1	0	1	1
4	0	1	0	0	1
5	0	1	1	1	0

Table 2: RI Centroid Distance

	1	2	3	4	5
1	0.00000	12.88647	19.06757	25.73029	26.40893
2	12.88647	0.00000	15.02550	13.13064	17.71784
3	19.06757	15.02550	0.00000	23.25883	9.55189
4	25.73029	13.13064	23.25883	0.00000	19.85757
5	26.40893	17.71784	9.55189	19.85757	0.00000

metric, and replace each non-zero term with either a zero or one depending on if the distance between the two points is great or less than that distance respectively. For example, in the case of Rhode Island, a distance metric of adjacency being county centroids no more than 30 miles apart renders all off diagonal elements 1, and a distance of less then 9 miles renders the whole matrix 0's. Preference for the distance metric over border adjacency are beliefs in local economic similarities being determined by distance between counties, versus simple boarder adjacency. Downsides are that even counties that are border adjacent may not be distance adjacent in the Midwest due to large county distances.

In this setting, if two points $(a, b) \in D_n$ share the same distance $d_\emptyset$ from a third county centroid $c \in D_n$, then c is adjacent for any preferred distance metric of adjacency $d \geq d_\emptyset$. I pick either the 25 mile to 50 mile adjacency points. I use this metric since entrepreneurs have empirically traveled moderate distances trying to find beneficial loan deals. In [Degryse and Ongena, 2005] sample of Belgian banks the max loan distance is 50 miles. Similarly [Agarwal and Hauswald, 2010] find that average bank applications come from about 10 miles away, with a standard deviation of about 21 miles, while accepted applications come from even closer to the bank (2.62 miles), with a smaller standard deviation (10.67 miles). Thus while most bank applications are local, applicants seem to be willing to drive moderate distances in search of favorable loan contracts.

Define the partition of the set $\mathcal{N}_n$ around the node i as (i, W_i, W_{-i}). The set W_i has cardinality equal to the sum of column (or row) i. The set W_{-i} contains all nodes not connected to i. Thus W_i is the set of neighbors. For each unit i, the partition (i, t, W_i, W_{-i}) defines the following partitions of the treatment and outcome vectors $(\underline{s_{it}}, \underline{\mathbf{s}_{W_i,t}}, \underline{\mathbf{s}_{W_{-i},t}})$ and $(Y_{it}, \mathbf{Y}_{W_it}, \mathbf{Y}_{W_{-i}t})$

Assumption 3.2 (No Multiple Versions of Treatment (Consistency))**.**

$$Y_{it} = Y_{it}(\mathbf{s}_{\mathbf{i}\mathcal{T}})$$

This assumption states that there are no additional treatments. Thus receiving different levels of CPP funds does not change the treatment level, or treatment status, of counties. This assumption implies that county i's potential outcome can depend on the whole history of all other counties treatment status up to that period. In my next assumption I assume the existence of a known link function that relates the treatment status of neighbors in each period to their influence on a subject county.

Assumption 3.3. *For each t, there exists a (possibly) vector valued function $g_{it} : \underline{s_{\min(t,2)}}^{W_i} \to \mathcal{G}_t$, $\forall i \in \mathcal{N}_n$, $\forall \underline{\mathbf{s_{W_i}}}, \underline{\mathbf{s_{W_{-i}}}}$, and $\forall \underline{\mathbf{s_{W_{it}}}}', \underline{\mathbf{s_{W_{-it}}}}' : g(\underline{\mathbf{s_{W_{it}}}}', \underline{\mathbf{s_{W_{-it}}}}') = g(\underline{\mathbf{s_{W_{it}}}}, \underline{\mathbf{s_{W_{-it}}}})$, the following equality holds,*

$$Y_{it}(\underline{s_{it}}, \underline{\mathbf{s_{W_{it}}}}, \underline{\mathbf{s_{W_{-it}}}}) = Y_{it}(\underline{s_{it}}, \underline{\mathbf{s_{W_{it}}}}', \underline{\mathbf{s_{W_{-it}}}}')$$

This assumption claims that given a node $i \in \mathcal{N}_n$, changes in which neighbor receive treatment, without changing the structure of treatment, does not impact node i's potential outcomes.

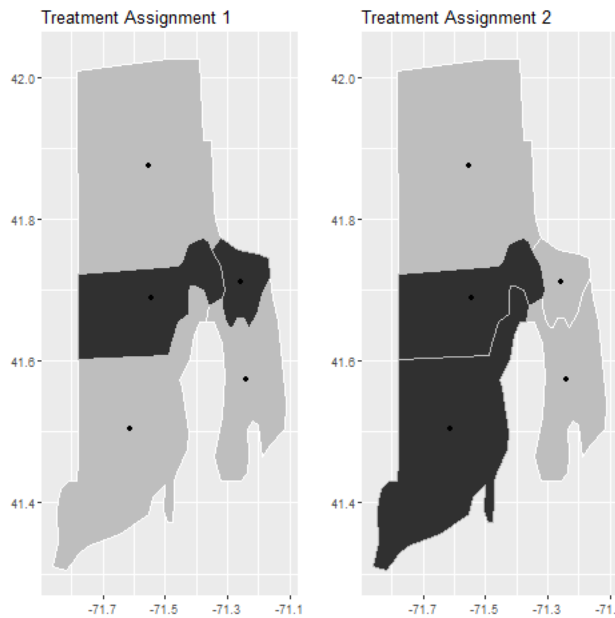
The link function $g(\cdot)$ is assumed to be a known function that relates the treatment status of neighbors to the potential outcome of a central node. For now I assume that it simply returns the number of treated individuals in a particular treatment category. Therefore, when researchers are dealing with a static problem, $t = \tau = 1$. $g(\cdot)$ is no longer vector valued, and simply returns a count for the number of treated neighbors a node i has. The assumption in this case states that the number of untreated individuals around specific node doesn't impact their outcomes. In the case of $t = 2$, $g(\cdot)$ now returns a count of the number of neighbors in the $(1, 0), (0, 1)$ and $(1, 1)$ treatment groups. Again, the number of non-treated $(0, 0)$ individuals around node i doesn't impact their outcome, nor does the treatment status of individuals outside their set of neighbors. Note that this structure rules out common interference effects like exponential decay.

Under the link function, "any neighbor treated as of period t" there are 16 potential outcomes. In each period, both $S_t \in \{0, 1\}$, and $G_t \in \{0, 1\}$. Under cases when $g(\cdot)$ is countable, there are $4 \times G^2$ potential outcomes. Returning to Rhode Island, assume that county "2" is my focus node in Figure 4. Under this setup, 2 is adjacent to every other county. Under assumption 3.3, the potential outcome for county 2 does not change. One implication of this is that if a node county is straddled on opposite ends by a heavily urbanized area, and on the other by a very rural area, regardless of which county receives treatment the potential outcomes for the node do not change. Assumption 3.3 might not be reasonable if the number of potential entrepreneurs looking for credit can exhaust the potential credit supply in a neighboring county.

Under this framework, each individual is subject to several treatments, both their own treatment grouping $\underline{s_{it}}$ as well as the grouping treatment G_{it}. The assignment mechanism is now the probability of the joint treatment in the whole sample, given all covariates and potential outcomes. Moreover, potential outcomes can be indexed by the individual treatment, and the neighborhood treatment function results. In period 1 this amounts to $2 \times \#G_{i1}$ combinations, and in period 2, $4 \times G_{i2}$ possible potential states. This potential outcome is defined only for a subset of nodes where G_i can take on value g_{it}.[3] Thus in

[3] Imagine the simple case where G_{it} counts the number of treated neighbors of any type in a given year. Thus for $g_{it} > m$, only counties with m or more neighbors can reach that potential state.

Figure 5: Treatment Assignment Example



period 1, individuals treatment status can be characterized as either $(0,g)$ or $(1,g)$ for $g \in \mathcal{G}_1$, or as $(0,0,g),(0,1,g),(1,0,g),(1,1,g)$ for $g \in \mathcal{G}_2$ in the second period.

I am interested in estimating the impacts of treatment on the treated. The total effect can be calculated as

$$\tau_t^{\underline{s_\tau}^k,\underline{s_\tau}^l}(\underline{s_\tau}^k) = \sum_{g\in\mathcal{G}} E(Y_t^{\underline{s_2^k}} \mid \underline{S_\tau} = \underline{s_2}^k, G_{it} = g) - E(Y_t^{\underline{s_2^j}} \mid \underline{S_\tau} = \underline{s_2}^k, G_{it} = 0)P(G_i = g) \tag{1}$$

$$t \geq T_0(\underline{s_\tau}^k), T_0(\underline{s_\tau}^l) \tag{2}$$

The leading term here is the potential outcome of an individual who receives the treatment chain $\underline{s_2^k}$ against the counter factual treatment chain $\underline{s_2^j}$, conditioned on having received $\underline{s_2^k}$. [Lechner, 2005, Lechner, 2008] show how to estimate this term using propensity scores under the SUTVA assumption. When this is relaxed, I need to decompose this into a *direct effect* and a *spillover effect.*

The *main effect* can be thought of as,

$$r_t^{\underline{s_2}^k,\underline{s_2}^l}(\underline{s_2}^k, g) = E[Y_{it}(\underline{S_{it}} = \underline{s_2}^k, G_{it} = g) - Y_{it}(\underline{S_{it}} = \underline{s_2^j}, G_{it} = g) \mid \underline{S_\tau} = \underline{s_2}^k, i \in V_g]$$

The last term $i \in V_g$ needs some explaining. The term V_g defines the set of nodes that can take on a particular values for $g(\cdot)$. This restricts the potential outcome of $Y(\cdot)$ to be comparable only to other nodes with similar number of potential values. For instance, in the case where G_i is the number of treated neighbors in the static case, V_g is the set of nodes with degree $W_i \geq g$, that is, with at least g neighbors. It is worth noting that each unit can belong to different subsets V_g depending on g. Under this framework, the average main effect is now

$$r_t^{\underline{s_2}^k,\underline{s_2}^l}(\underline{s_2}^k) = \sum_{g\in\mathcal{G}} r(g)^t P(G_{it} = g)$$

The spillover effect can be calculated as

$$\upsilon_t^{\underline{s_2}^k,\underline{s_2}^l}(g) = E[Y_{it}(\underline{S_{it}} = \underline{s_2}^k, G_{it} = g) - Y_{it}(\underline{S_{it}} = \underline{s_2}^k, G_{it} = 0) \mid \underline{S_\tau} = \underline{s_2}^k, i \in V_g]$$

where again, the average spillover effect can be calculated as,

$$\upsilon_t^{\underline{s_2}^k,\underline{s_2}^l} = \sum_{g\in\mathcal{G}} \upsilon(g)^t P(G_{it} = g)$$

To help us estimate this effect, I assume the following ignorability conditions.

Assumption 3.4. *For almost every $x \in \mathbb{X}$, where $\mathbb{X}$ is the support of X*

1. *S_1, G_1 is independent of $Y_2((0,0)), Y_2((0,1)), Y_2((1,0)), Y_2((1,1))$ conditional on $X = x_0, G_{i0} = g$*

2. *S_2, G_2 is independent of $Y_2((0,0)), Y_2((0,1)), Y_2((1,0)), Y_2((1,1))$ conditional on $X = x_1, S_1 = s, G_1 = g, X = x_0, G_{i0} = g$*

3. *$\eta < Pr(S_{i,1} = 1, G_{i,1} = g_1 \mid X_{0,i} = x_0, X_{0,-i} = x_{0,-i}) < 1-\eta$, $\eta < Pr(S_2 = 1, G_{i,2} = g_2 \mid X_{1,i} = x_{1,i},\ X_{0,i} = x_{0,i},\ S_{i,1} = s, G_{i,1} = g,\ X_{1,-i} = x_1',\ X_{0,-i} = x_{0,-i},\ S_{1,-i} = s_{-i}) < 1-\eta$ for some $\eta > 0$*

The first part of this assumption states that all four potential outcomes are independent of treatment in period 1 conditional on period 0 covariates, and the number of treated neighbors in period 0 (which, by design is always 0). Similarly, This is standard in the treatment effects literature. The second assumption differentiates this problem from the standard multiple treatments literature in allowing treatment in the second period to be independent of next period potential outcomes conditional on X *and* the previous period's decision to be treated.

The propensity score is now defined to be the probability of own treatment status and number of treated neighbors in period 1 conditional on period 0 covariates for the subject, and their neighbors. In period 2 this is repeated, but now contains period 1 covariates for both subject and their neighbors, including the complete vector of period 1 considerations.

I now wish to leverage these assumptions into two different estimation strategies. The first I discuss is a matching estimator. This estimator leverages the propensity scores of counties being in different treated categories, and their transition probabilities, to match counties based on pre-treatment characteristics. Versions of this estimator in non-dynamic settings have previously been worked out by [Liu et al.,]. Similar papers include [Sobel, 2006, Hudgens and Halloran, 2008, Tchetgen Tchetgen and VanderWeele, 2012].

3.1 Matching Estimator

The first method for estimating Equation (1) I employ is a matching estimator. Throughout I take the link function "any neighbor having received a treatment" as given. This simplifies some of the analysis. Most importantly, as long as a county has any neighbors, they never get dropped from being a feasible match. Under more complex link functions researchers need to track the feasability of a county to be a viable match.[4] No county in my sample has no neighbors, and thus this concern is dropped from the estimators. Since my set of treatments is no longer binary, this introduces new notation. Define the term

$$D_{A,t,i} = I\{\underline{s_{it}}^k = A\}$$

That is, if in period t, individual i's treatment status is equal to a particular target sequence. Equivalently, terms like $N_{A,t}$ denote the number of individuals in a particular target sequence in period t. Further subscripts, such as $N_{A,t,g}$ denote the number of individuals in period t with treatment sequence A and number of treated neighbors g. I also assume the link function $g(\cdot)$ takes on only two values, 1 if the subject has any treated neighbor in period t, and 0 otherwise. I can generate the three estimators for the total effect, main effect, and spill over effect as,

$$\hat{\tau}_t^{\underline{s_2}^k,\underline{s_2}^l}(\underline{s_2}^k) = \frac{1}{N_{\underline{s_2}^k},t} \sum_{D_{\underline{s_2}^k,t,i}=1} (Y_{it} - \hat{Y}_{it}(\underline{s_2}^l, G_{it}=0)) \tag{3}$$

$$\hat{r}_t^{\underline{s_2}^k,\underline{s_2}^l}(\underline{s_2}^k, g) = \frac{1}{N_{\underline{s_2}^k,t,g}} \sum_{D_{\underline{s_2}^k,t,i}=1, G_{it}=g} (Y_{it}(G_{it}=g) - \hat{Y}_{it}(\underline{s_2}^l, G_{it}=g)) \tag{4}$$

$$\hat{\upsilon}_t^{\underline{s_2}^k,\underline{s_2}^l}(g) = \frac{1}{N_{\underline{s_2}^l,t,g} + N_{\underline{s_2}^k,t,g}} \sum_{i,j:D_{\underline{s_2}^j,t,i}=1, G_{it}=g} (Y_{it}(D_{\underline{s_2}^j,t,i}=1, G_{it}=g) - \hat{Y}_{it}(D_{\underline{s_2}^j,t,i}=1, G_{it}=0)) \tag{5}$$

Here the counterfactuals are constructed such that

$$\hat{Y}_{it}(\underline{s_2}^l, G_{it}=0) = \begin{cases} 0 & D_{\underline{s_2}^l,t,i}=1, \text{ or } D_{\underline{s_2}^k,t,i} \neq 1 \text{ or } G_{it} \neq 0 \\ \frac{1}{M}\sum_{j \in \mathbb{J}_M(i,t,G_{it}=0,D_{\underline{s_2}^l,t,i}=1)} Y_{jt} & D_{\underline{s_2}^k,t,i}=1 \end{cases}$$

[4] Under another common link function, "number of neighbors treated", this would require matching only on the subset of counties with at least G_i number of neighbors as well as the dynamic propensity scores.

$$\hat{Y}_{it}(\underline{s_2}^l, G_{it} = g) = \begin{cases} 0 & D_{\underline{s_2}^l,t,i} = 1, \text{ or } D_{\underline{s_2}^k,t,i} \neq 1 \text{ or } G_{it} \neq g \\ \frac{1}{M} \sum_{j \in \mathbb{J}_M(i,t,G_{it}=g,D_{\underline{s_2}^l,t,i}=1)} Y_{jt} & D_{\underline{s_2}^k,t,i} = 1 \end{cases}$$

$$\hat{Y}_{it}(D_{\underline{s_2}^j,t,i} = 1, G_{it} = 0) = \begin{cases} 0 & D_{\underline{s_2}^j,t,i} = 1 \neq 1 \text{ or } G_{it} \neq 0 \\ \frac{1}{M} \sum_{j \in \mathbb{J}_M(i,t,G_{it}=0,D_{\underline{s_2}^j,t,i}=1)} Y_{jt} & D_{\underline{s_2}^j,t,i} = 1 \end{cases}$$

Where $\mathbb{J}_M(i,t,G_{it},D_{\underline{s_2}^k,t,i} = 1)$ are M closest counties by propensity score in period t condition on a specific $G_{it} = g$ and $D_{\underline{s_2}^k,t,i} = 1$. Unit j is the M-th closet match to observation j if, $D_{\underline{s_2}^l,t,i} = 1$. In period 1 I use the estimates for counties have a specific treatment status x link function value

$$P(S_{i,1} = s_1^k, G_{i1} = g_1 \mid X_{0,i} = x_0, X_{0,-i} = x_{0,-i})$$

That would be equivalent to the regular matching on a propensity score. For $t > 1$ I use both the stage 1 propensity score, along with the stage 2 propensity score for $S_2^k = s_2^k \mid S_1^k = S_1^k$, which under Assumption 3.4 is the joint probability.

$$Pr(S_{i,2}^k = s_2^k, G_{i,2} = g_2 \mid X_{1,i} = x_{1,i},\ X_{0,i} = x_{0,i},\ S_{i1} = s_1, G_{i1} = g_1, X_{1,-i} = x_1',\ X_{0,-i} = x_{0,-i},\ S_{1,-i} = s_{-i})$$

In later periods all counties receive the same (no) treatment, so it remains sufficient to match on those two propensity scores for $t = 3, 4, \ldots$. Algorithms for calculating the total effect are in Table 11, the main effect in Table 12, and the indirect effect in Table 13.

3.2 Synthetic Control

The above estimators have several desirable properties, including not requiring a functional form, and under a strong belief in Assumption 3.4 decent ability to recover main, interference, and total average treatment on the treated effects. However, they still require estimating the propensity score, and such approximations carry natural drawbacks. My alternative specification is to utilize synthetic control methods, where I assume a linear data generating process with unknown factor-loading components. Matching is then carried out on a longer time period of pre-treatment covariates to estimate the forward looking ATT's. In this section I show that the objective functions for the classical synthetic control estimator can be augmented in order to estimate the Total Effect, Main Effect, and Interference Effect under 3.3. These estimates differ from the traditional synthetic control estimator which was for a single individual ([Abadie et al., 2010]), and instead we impose an ex ante assumption that these treatments were the same for all individuals in each treated period. This allows multiple estimation methods, including both the methods outlined below, and by recent work by [Cao and Dowd, 2019].

Similar to the matching estimators I want to extend synthetic control methods to estimate the direct, spillover, and total effect. Under my assumptions, I estimate the group mean effect for both the main and interference effect. It is then possible to construct the total effect estimator using empirical estimates for the probability of each group assignment path. This is carried out by conditioning on a perfect set of matches allows the creation of simple estimates for the main effect and interference effect. This is done by restructuring what makes up "non-treated" status as to generate appropriate counterfactuals over counties potential outcomes. This motivates an estimator for the main effect as,

$$\begin{aligned} Y_{it}^{T,\underline{g},MAIN}(\underline{s_{\mathcal{T}}}^l) &= Y_{it}^{N,\underline{g},MAIN} + r_t(\underline{s_{\mathcal{T}}}^l) D_{\underline{s_{\mathcal{T}}}^l,i,t} \\ Y_{it}^{N,\underline{g},MAIN} &= \delta_t + \theta_t Z_i + \lambda_t \mu_i + \upsilon_t(\underline{g}) + \epsilon_{it} \end{aligned}$$

where δ_t is an unknown common factor with constant factor loadings across units, Z_i is a $(k \times 1)$ vector of observed covariates (not affected by the intervention) that do not vary across times, θ_t $(1 \times k)$ vector of unknown parameters, λ_t is a $(1 \times F)$ vector of unobserved common factors, μ_i is an $(F \times 1)$ vector of unknown factor loadings, and the error terms ϵ_{it} are unobserved transitory shocks at the region level with zero mean. Under my link function, $\underline{g}$ is a set of 4 binary variables, and thus υ_t is a 1×4 vector of coefficients in period t for a given neighbor treatment status. For each estimator I view T_0 as the first date of treatment. $r_t(\underline{s_{\mathcal{T}}}^l)$ is the main effect for counties in the $\underline{s_{\mathcal{T}}}^l$ treatment group. For all treatment groups, the counterfactuals is assumed to be the same. This motivates the estimator,

$$r_t(\underline{s_{\mathcal{T}}}^l) = \frac{1}{N_{\underline{s_{\mathcal{T}}}^l}} \sum_{\underline{g} \in \underline{\mathcal{G}_t}} \left(Y_{it}^{T,\underline{g},MAIN}(\underline{s_{\mathcal{T}}}^l) - Y_{it}^{N,\underline{g},MAIN} \right) \tag{6}$$

Similar to the matching estimator, I treat $Y_{it}^{T,g,MAIN}(\underline{s_{\mathcal{T}}}^l)$ as observed, and wish to create a composite synthetic county that behaves as if $Y_{it}^{N,g,MAIN}$ would. Estimation of this term now follows exactly from previous synthetic control estimators ([Abadie et al., 2010]).

The interference effect can equivalently be defined any $g \in \mathcal{G}_t$,

$$\begin{aligned} Y_{it}^{k,\underline{g},SPILL} &= Y_{it}^{N} + \upsilon_t(\underline{g}) \\ Y_{it}^{k,(0,0),SPILL} &= \delta_t + \theta_t Z_i + \lambda_t \mu_i + r_t I\{k=1\} + \epsilon_{it} \end{aligned}$$

Where $k \in \{0,1\}$ as a proxy for whether or not a county is treated or not. This potential outcome structure is equivalent to comparing positive neighbor treatment status against the no-neighbor treatment status, conditioned on the individuals own treatment status in period t. It leads to the similar infeasible estimator for the Average Treatment Spillover Effect for the Treated (ATST) as,

$$\hat{\upsilon}_t^S(g,0) = \frac{1}{N_T}\sum_{i=1}^{N_T}\left(Y_{it}^{k,g,SPILL} - Y_{it}^{k,0,SPILL}\right) \tag{7}$$

As above, I treat $Y_{it}^{k,g,SPILL}$ as observed, and use synthetic control methods to create a composite county that behaves as $Y_{it}^{k,0,SPILL}$ does by matching along covariates in the pre-treatment time periods. Finally, from Equation (1), the total effect can be estimated as,

$$\hat{\tau}_{it}(\underline{s_\tau}^l) = \hat{r}_{it}(\underline{s_\tau}^l) + \sum_{g=0}^{1}\hat{\upsilon}_t^S(g,0)\hat{P}(g) \tag{8}$$

3.3 Sub Sampling

The following procedure is used to generate an independent sample of observations assuming one-adjacency independence given a graph (D_n, W_n). That is, for any i, j, if $j \notin W_i$, then Y_{it} and Y_{jt} are independent of each other. Under this assumption, I created the following algorithm to induce independence on the graph.

1. Create an empty set $I = \{\emptyset\}$.
2. Given the graph (D_n, W_n) split it up into treated and untreated subgraphs (D_n^T, W_n^T), (D_n^N, W_n^N)
3. Chose $i \in D_n^T$, add to I. Then remove all $j \in D_n$ such that $j \in W_i$. This includes both from the treated and untreated subgraphs.
4. Generate the resulting sub-graph (D_n', W_n').
5. Repeat until $D_n' = \{\emptyset\}$

Using this method I am able to induce independence on my sample and use straightforward applications of Inverse Propensity Score Weighting (see [Chan et al., 2016]) or matching estimators ([Abadie and Imbens, 2006]). This assumes there is no interference effect, and instead the concern is in removing spatial correlation between closely linked treated and untreated counties. This process has clear downsides. First, treated nodes with no other treated neighbors appear in the independent sample with probability 1. Dissecting the relative probability that members of "chains" of treated individuals appear can also become quite complex. This biases bootstrap style estimators that try to use this sub-sampling procedure. Secondly, the assumptions used to generate the procedure do not account for spill overs, and only solve issues with spatial correlation in errors under a spatial correlation framework of a random field (Conley (1999)), versus more traditional spatial AR techniques (LeSage (1999), Kelejian and Prucha (2007)). However, these results provide benchmark solutions for a total effect under the assumption of a random field, or the main effect under the presence of no interference effect.

4 Data and Preliminary Analysis

My preliminary dependent variable of interest is county level firm entry rates. I use information on new firm entrants from the Census Statistics of US Businesses & Business Information Tracking Series.[5]. The data characterizes both the total number of establishments, the change in establishments, the number of births (new entrants), and deaths (exits). The data set tracks this information from 1999 to 2015. A number of states report zero firm entry over my sample. So I add a small incremental number to rule out negative infinite numbers.

I then generate aggregated CPP purchases at the county-year level using bank-city-date transaction data provided by the Treasury Department. I then create an indicator value on whether or not a county received any amount at all to indicate treatment status. That is

$$Treated_{i,t} = 1\{CPP_{i,t} > 0\}$$

This again is motivated by the hard limits on the amount of funds the Treasury Department could provide to participating banks, and the low number of banks that received CPP funds in each county.[6]

Several different time indices are used. I use both "treated in 2008" and "treated in 2009" as natural indexes. I also utilize a second pair of treatment statuses "treated in the first 9 months after program starts" or "treated between 9-18 months after program starts." Additionally I leverage county level unemployment rates and the size of the labor force using the BLS's Local Area Unemployment statistics,[7].

I further compile mean bank characteristics at the county level from FDIC call sheet data,[8] Following [Lei, 2013] I include tier 1 ratio, troubled assets ratio, annualized Return on Assets, Cash-to-Assets ratio, and loan-to-deposits ratio.[9] These proxy for local community bank health that the Federal Regulators may have observed when deciding which banks to accept into the CPP program.

Finally, I have a set of instruments that explain mean bank political connections in a county. These instruments include whether or not banks had a board member serving as a local Federal Reserve director, if the local house representative was on the Subcommittee for Finance and Insurance, local donations from Finance, Insurance, and Real Estate as a percent of the representatives total donations, and whether or not

[5] `https://www2.census.gov/programs-surveys/susb/`

[6] Traditional analysis has removed the 20 largest banks from the analysis (see [Lei, 2013]). In turn this would imply removing the counties with the 20 largest counties. The list of banks include, Goldman Sachs, J.P.Morgan Chase Bank, Keybank (Keycorp), PNC Bank, Fifth Third Bank, Bank of America, BB&T Bank (BB&T Corp), State Street, U.S. Bank (U.S Bancorp), Wells Fargo Bank, Suntrust Bank, Citibank, Capital One, Regions Bank, Bank of New York Mellon, Northern Trust Company, Comerica Bank, M&T Bank, Marshall&Ilsley Bank, and Morgan Stanley. In practice this excludes New York, NY; Charlotte, NC; Boston, MA; Minneapolis, MN; Cleveland, OH; Pittsburgh, PA; Cincinnati, OH; Atlanta, GA; McLean, VA; Birminham, AL; Chicago, IL; Dallas, TX; Buffalo, NU; and Milwauke, WI. However, my estimation method allows us to keep some of the information these counties have in my data set, though I do not use them in my treated sample.

[7] `https://www.bls.gov/lau/`

[8] `https://www.fdic.gov/regulations/resources/call/index.html`

[9] I calculate these values directly from call sheet data from 2008Q3. Tier 1 Ratio is calculated directly in the Call Sheets as RCON7206. Troubled Asset Ratio is loans 90 days past due/total capital. For loans I combined 90 Days Past Due C&I Loans (RCON5460) and All Other Loans Past Due 90 Days or MOre (RCON5460). For Total Capital I took Total Assets (RCON2170) and subtracted Total Liabilities (RCON2948). Return on Assets was Net Income (RIAD4340) divided by Total Assets. Cash to Assets was Cash and Due From Depositors (RCON0010) divided by Total Assets. Loan to Deposits Ratio was Loans, Leases, Net Unearned Income (RCONB528) divided by Total Deposits (RCON2200).

the local house representative was a democrat in 2008 Q3.

To generate my county adjacency matrix, I use the Census County Adjacency File[10] The County Adjacency File uses a strict definition of counties that neighbor each other, but includes counties whose border lies in a body of water as neighbors. Unweighted information on county averages are covered in tables 7-10.

Comparable estimates with the Synthetic Control methods use number of neighbors treated, unemployment rate, (county mean) troubled asset ratio, return on assets, loans to deposits, labor force, mean neighbor characteristics for unemployment rate, troubled asset ratio, return on assets, loans to deposites, and labor force. This controls for the larger regional characteristics that might have impacted link function assignment.

5 Results

5.1 Matching Estimator Results

Under the link function "any neighbor having received a treatment" there are 16 potential outcomes, and conditioned on a particular treatment sequence $\underline{s_2}$ 4 potential outcomes based around link function outcomes in each time period. This generates a single main effect, and 3 relevant spillover effects.[11]

5.2 Synthetic Control Results

I currently report just the main effect estimates. These are carried out using the Microsynth package in R. My set of time invariant variables in a null set. My output variables is just log births, and I minimize the match along number of neighbors treated, unemployment rate, (county mean) troubled asset ratio, return on assets, loans to deposits, labor force, mean neighbor characteristics for unemployment rate, troubled asset ratio, return on assets, loans to deposites, and labor force. By forcing the match on number of neighbors treated to be as close to zero as possible acts as matching on the interference effect, and generates estimates for the direct response for average treatment on the treated.

100 permutation draws are drawn from the untreated counties to construct an exact test for the existence of a true treatment effect. These draws are represented by the gray lines, while the red lines represent the path for the estimates for each periods Average Treatment for the Treated represented by net firm entry. Across all three estimated main effects the data rejects the existence of a treatment effect at the 10% level.

[10] `https://www.census.gov/geo/reference/county-adjacency.html`. I plan to do further analysis using the NBER database on County Distance Database. `http://www.nber.org/data/county-distance-database.html`

[11] $(0,1)$, $(1,0)$, and $(1,1)$ relative to $(0,0)$.

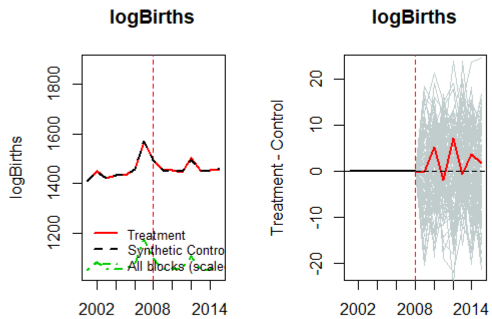


Figure 6: Main Effect for (1,0) Counties

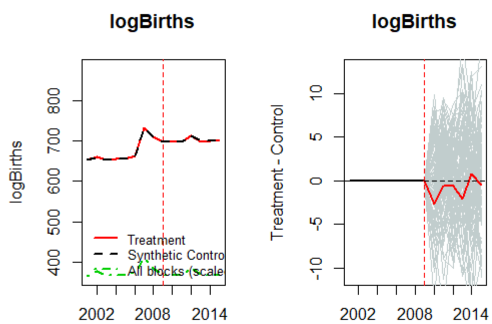


Figure 7: Main Effect for (0,1) Counties

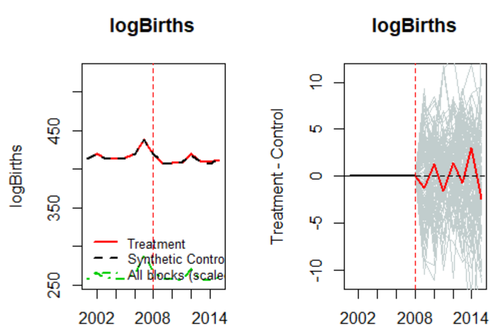


Figure 8: Main Effect for (1,1) Counties

5.3 Sub Sampling Results

Current models are estimated using Change in Home Price Index, Unemployment Rate, Total Population, Share Non-White, Average Loan to Deposits Ratio, Average Bank Return on Investment, If one bank has a board member serving as a regional Federal Reserve Director, and if the local House of Representatives member is on the Finance and Insurance Subcommittee. This is a known bad practice ([Wooldridge, 2009]), and I aim to tidy this up to include an instruments version of the treatment based around an estimated propensity score. For for each group, I estimate both a Propensity Weighting estimator ([Chan et al., 2016]), and a matching estimator ([Abadie and Imbens, 2006]).

I estimate the Average Treatment on the Treated in net firm entry[12] for counties who received treatment in the first 7 months but not in the subsequent time period, in 2008 but not 2009, in 2009 but not 2008. For each I then estimate an 'extended treatment zone' model, where I assume every county around them was also in an effective treatment zone. In all cases, I then remove counties immediately adjacent to all treated counties from the set of plausible untreated counties to match with.[13] I do this because there may be small but residual spill over between treated counties and those immediately around them.

The ATT can be viewed as one step ahead impacts of Treatment. This gives us four periods for counties either treated in the first 7 months following November 14th, 2008, or treated before Jan 1st, 2009, and three periods for counties treated between Jan 1st, 2009 and Dec 31st, 2009. For each model class I report the Average Treatment on the Treated, the Standard Error, and the resulting P value.

Table 3: ATT One Step Ahead

Name	ATT	SE	p.val
Interval Prop	0.0027353	0.00170278	0.1082
Interval Match	0.0017795	0.0021697	0.41212
Interval Ext Prop	0.0013587	0.0018306	0.4579
Interval Ext Match	0.0027473	0.022921	0.23069
Treated 08 Prop	0.00084703	0.00247123	0.7318
Treated 08 Match	-0.001557	0.0026545	0.5575
Treated 08 Ext Prop	0.0024647	0.0026511	0.3525
Treated 08 Ext Match	0.0020552	0.0028978	0.47817
Treated 09 Prop	-0.0007873	0.0014115	0.577
Treated 09 Match	-0.00018906	0.0018548	0.91881
Treated 09 Ext Prop	-0.00222590	0.00136390	0.102
treated 09 Ex Match	-0.0012233	0.001552	0.43055

6 Conclusion

In this paper I examine the role that counties having banks that received CPP funds had on subsequent firm entry. I utilize a mix of matching, synthetic control, and sub sampling to try and parse out both

[12] This was calculated using the Census County Business Patterns, before I had access to the SUSB data set.

[13] Note, this implies that all new counties added between the initial treated and extended treatment range are not included in the first estimation strategy as counties viewed as either untreated or treated.

Table 4: ATT two Step Ahead

Name	ATT	SE	p.val
Interval Prop	-0.00124286	0.00141347	0.3792
Interval Match	-0.00069937	0.0019654	0.72196
Interval Ext Prop	-0.00203849	0.00177112	0.2497484
Interval Ext Match	-0.00038661	0.0020675	0.85166
Treated 08 Prop	-0.0022506	0.0019049	0.2374
Treated 08 Match	-0.00082901	0.0025923	0.74913
Treated 08 Ext Prop	-0.00240407	0.00237312	0.3110395
Treated 08 Ext Match	-0.0019562	0.0029416	0.50604
Treated 09 Prop	0.00226397	0.00143733	0.1152
Treated 09 Match	0.0018394	0.0019326	0.34122
Treated 09 Ext Prop	0.00111247	0.00127751	0.3839
treated 09 Ex Match	-0.0008781	0.0014535	0.54576

Table 5: ATT three Step Ahead

Name	ATT	SE	p.val
Interval Prop	0.00089019	0.00153397	0.5617
Interval Match	0.0031404	0.0020697	0.12919
Interval Ext Prop	-0.00053458	0.00159921	0.7382
Interval Ext Match	0.00094758	0.0020834	0.64924
Treated 08 Prop	0.0014894	0.0024949	0.5505
Treated 08 Match	0.00096561	0.0023419	0.6801
Treated 08 Ext Prop	0.0060254	0.0022129	0.006472
Treated 08 Ext Match	0.0074124	0.0027348	0.0067204
Treated 09 Prop	-0.00243489	0.00169923	0.1519
Treated 09 Match	-0.0037673	0.0021096	0.074141
Treated 09 Ext Prop	-0.00343415	0.00151520	0.02342
treated 09 Ex Match	-0.0023231	0.0017984	0.19646

Table 6: ATT Four Steps Ahead

Name	ATT	SE	p.val
Interval Prop	-3.5330e-03	1.7614e-03	0.04488
Interval Match	-0.0025952	0.0022978	0.25872
Interval Ext Prop	-0.00553142	0.00192987	0.004154
Interval Ext Match	-0.0048049	0.0023976	0.045065
Treated 08 Prop	-0.0029346	0.0021983	0.18189
Treated 08 Match	-0.0036463	0.0033968	0.28307
Treated 08 Ext Prop	-0.0101920	0.0040789	0.0124656
Treated 08 Ext Match	-0.0061285	0.0044011	0.16377

the direct effect of a county receiving CPP cash, and the spillover effect of neighboring counties receiving CPP injections. Examining firm entry has several benefits over direct bank level responses. Particularly among small firms, relationship lending is often contextual to intra-bank structure, and formally modeling this might be difficult. Estimation might be biased by poor understanding of this mechanism. However, higher firm entry was still a preferred outcome of policy makers at this time as a way of encouraging new job growth and aiding recovery efforts.

Despite the large outlay of cash from the Federal government, across a variety of specifications there is mixed at best evidence for counties receiving TPP funds having higher firm entry. In fact, the strongest evidence is that the Four Step Ahead impacts were negative relative to untreated counties. In many of the other periods, model estimates are often contradictory of each other. While the one step ahead forecast has no ATT that has p-value less than .1, two are close, the 7 month interval estimate and Treated in 2009 with extended borders, but the resulting point estimates are opposite signs with almost equal magnitude. The two step ahead ATT is uniformly unidentified. The three step ahead ATT again has several results with positive and statistically significant results, and several with negative and statistically significant results. Thus, I argue counties receiving CPP funds did not impact future county business entry.

Care still needs to be taken when talking about the limits of these results. Many other programs were investing money into these counties simultaneously, including other TARP programs, that administrative files currently don't give us information to add to my analysis. However, CPP funds still made up a large share of total outlays by the government between late 2008 and early 2009, with little sign of improvement in business conditions among participating counties.

References

[Abadie et al., 2010] Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.

[Abadie and Imbens, 2006] Abadie, A. and Imbens, G. (2006). large sample properties of matching estimators for average treatment effects. *econometrica*, 74(1):235–267.

[Agarwal and Hauswald, 2010] Agarwal, S. and Hauswald, R. (2010). Distance and private information in lending. *the review of financial studies*, 23(7):2757–2788.

[Bassett, 2016] Bassett, W F; Demiralp, S. L. N. (2016). Government support of banks and bank lending. *Working Paper*.

[Berger and Udell, 2002] Berger, A. N. and Udell, G. F. (2002). Small business credit availability and relationship lending: The importance of bank organisational structure. *The Economic Journal*, 112(477):F32–F53.

[Blau et al., 2013] Blau, B. M., Brough, T. J., and Thomas, D. (2013). Corporate lobbying, political connections, and the bailout of banks. *Journal of banking and finance*, 37(8):3007–3017.

[Cao and Dowd, 2019] Cao, J. and Dowd, C. (2019). Inference in synthetic controls with spillover effects. *Working Paper*.

[Chan et al., 2016] Chan, K., Yam, S., and Z, Z. (2016). globally efficient nonparametric inference on average treatment effects by empirical balancing calibration weighting. *jounral of the RSS statistical methodology series B*, 78(3):673–700.

[Clementi and Palazzo, 2016] Clementi, G. L. and Palazzo, B. (2016). Entry, exit, firm dynamics, and aggregate fluctuations. *American economic journal: macroeconomics*, 8(3):1–41.

[Cole, 2012] Cole, R. A. (2012). How did the financial crisis affect small business lending in the united states? *Small business administration.*

[Contessi and Francis, 2011] Contessi, S. and Francis, J. L. (2011). Tarp beneficiaries and their lending patterns during the financial crisis. *Federal Reserve Bank of St. Louis Review*, 93(3):105–125.

[Degryse and Ongena, 2005] Degryse, H. and Ongena, S. (2005). Distance, lending relationships, and competition. *jouranl of finance*, 60(1):231–266.

[Duchin and Sosyura, 2014] Duchin, R. and Sosyura, D. (2014). safer ratios, riskier portfolios: banks' response to government aid. *journal of financial economics*, 113(1):1–28.

[Forastiere et al., 2018] Forastiere, L., Airoldi, E., and Mealli, F. (2018). Identification and estimation of treatment and interference effects in observational studies on networks. *arXiv Working Paper 1609.06245v4*.

[Gourio et al.,] Gourio, F., T, M., and M, S. *American economic review.*

[Hudgens and Halloran, 2008] Hudgens, M. G. and Halloran, M. E. (2008). Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–842.

[Imai and Ratkovic, 2015] Imai, K. and Ratkovic, M. (2015). Robust estimation of inverse probability weights for marginal structural models. *Journal of the American Statistical Association*, 110(511):1013–1023.

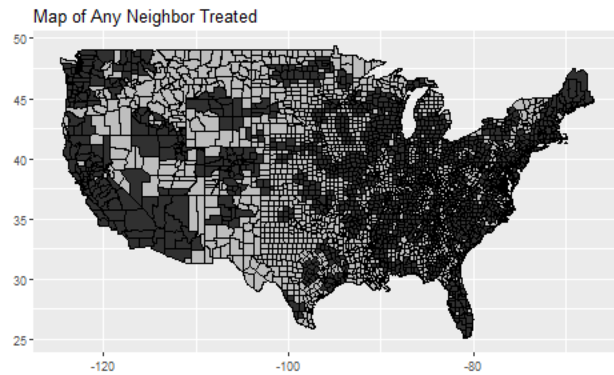
[Lechner, 2005] Lechner, M. (2005). Sequential matching estimation of dynamic causal models. *University of St. Gallen, Discussion paper.*

[Lechner, 2008] Lechner, M. (2008). matching estimation of dynamic treatment models: some practical issues. *advances in econometrics, modelling and evaluating treatment effects in econometrics*, 21:289–333.

[Lee and Mukoyama, 2015] Lee, Y. and Mukoyama, T. (2015). Productivity and employment dynamics of us manufacturing plants. *Economics letters*, 136:190–193).

[Lei, 2013] Lei, L. (2013). Tarp funds distribution and bank loan supply. *Journal of banking & finance*, 37(12):4777–4792.

Figure 9: Map of All Treated Counties & Neighbors



[Liu et al.,] Liu, L., Hudgens, M., and Becker-Dreps, S. On inverse probability-weighted estimators in the presence of interference.

[Robins et al., 2000] Robins, J., Hernán, M., and Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5):550–560.

[Sobel, 2006] Sobel, M. (2006). hat do randomized studies of housing mobility demonstrate? causal inference in the face of interference. *Journal of the American Statistical Association*, 101:1398–1407.

[Tchetgen Tchetgen and VanderWeele, 2012] Tchetgen Tchetgen, E. and VanderWeele, T. (2012). On causal inference in the presence of interference. *Statistical Methods in Medical Research*, 21:55–75.

[Veronesi and Zingales, 2010] Veronesi, P. and Zingales, L. (2010). Paulson's gift. *journal of financial economics*, 97:339–368.

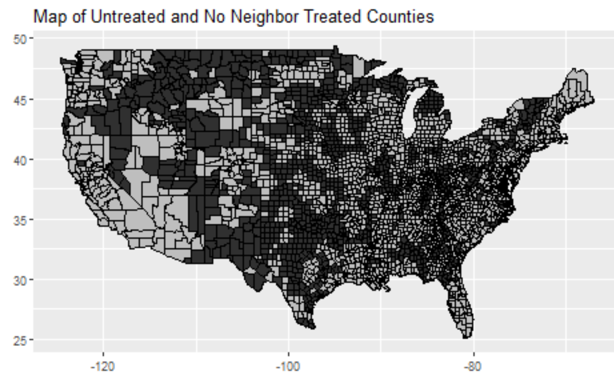
[Wooldridge, 2009] Wooldridge, J. (2009). Should instrumental variables be used as matching variables? *Working Paper*.

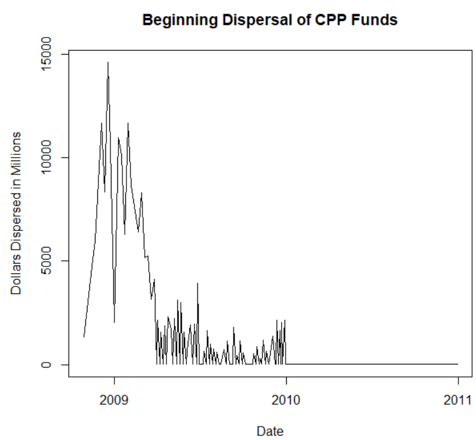
7 Appendix

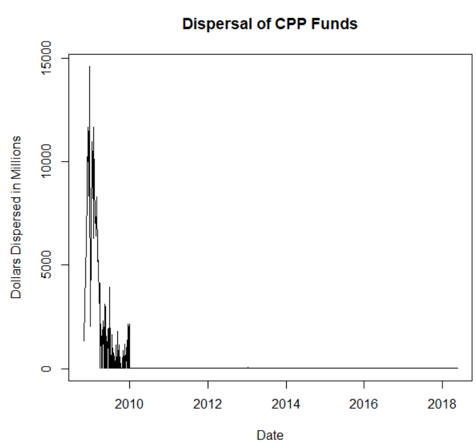
7.1 Figures

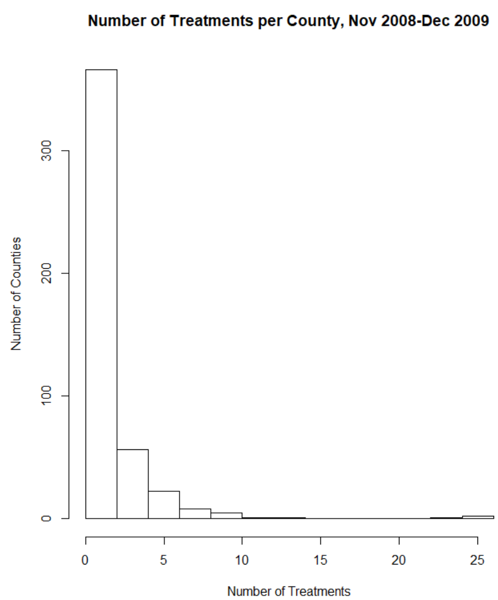
7.2 Summary Statistics

Figure 10: Map of all Counties that Received No Treatment nor did a neighbor









Number of Treatments per County, Nov 2008-Dec 2009
Number of Counties
0
100
200
300
0
5
10
15
20
25
Number of Treatments

Table 7: Untreated In All Periods Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
births2009	1,669	−0.024	0.027	−0.143	−0.038	−0.010	0.347
births2010	1,669	−0.006	0.026	−0.402	−0.018	0.006	0.203
births2011	1,669	−0.015	0.024	−0.158	−0.028	−0.003	0.142
births2012	1,669	0.004	0.033	−0.530	−0.012	0.019	0.160
birthsNoCons2009	1,669	−0.019	0.027	−0.154	−0.032	−0.005	0.360
birthsNoCons2010	1,669	−0.002	0.027	−0.442	−0.014	0.011	0.213
birthsNoCons2011	1,669	−0.012	0.025	−0.168	−0.025	0.0005	0.141
birthsNoCons2012	1,669	0.005	0.034	−0.537	−0.011	0.022	0.150
manufBirths09	1,669	−0.048	0.097	−0.693	−0.091	0.000	0.693
manufBirths10	1,669	−0.025	0.097	−0.693	−0.065	0.004	0.693
manufBirths11	1,669	−0.020	0.099	−1.099	−0.057	0.016	0.693
manufBirths12	1,669	0.006	0.117	−1.099	−0.042	0.053	0.693
change.index.2008	1,669	−0.027	4.645	−35.550	−1.840	2.260	20.630
Subcomm..on.FI	1,669	0.056	0.229	0	0	0	1
Local.FIRE.donation	1,669	0.020	0.020	0.000	0.005	0.029	0.127
Dem.	1,669	0.055	0.227	0	0	0	1
Treated.2009	1,669	0.000	0.000	0	0	0	0
Treated.2013	1,669	0.000	0.000	0	0	0	0
Treated.2008	1,669	0.000	0.000	0	0	0	0
Treated.2012	1,669	0.000	0.000	0	0	0	0
Unem.Rate	1,669	5.796	1.948	1.700	4.400	6.900	22.600
tier1ratio	1,669	1,628.103	793.686	9	1,014.5	2,199.7	3,519
troubledassetratio	1,669	407.038	512.780	1	1	688.3	2,027
ROA	1,669	2,767.418	1,256.301	7	1,918	3,774	5,141
cash.to.assets	1,669	2,609.905	1,215.100	34	1,686	3,515.3	5,131
loan.to.deposits	1,669	2,260.653	1,139.609	13	1,387	3,059.7	4,792
share_h	1,669	0.070	0.116	0.004	0.015	0.064	0.960
share_b	1,669	0.081	0.131	0.000	0.005	0.089	0.835
share_aa	1,669	0.009	0.015	0.0003	0.003	0.009	0.234
treatFst.ex	1,669	0.000	0.000	0	0	0	0

Table 8: Treated 2008 Summary Statistics

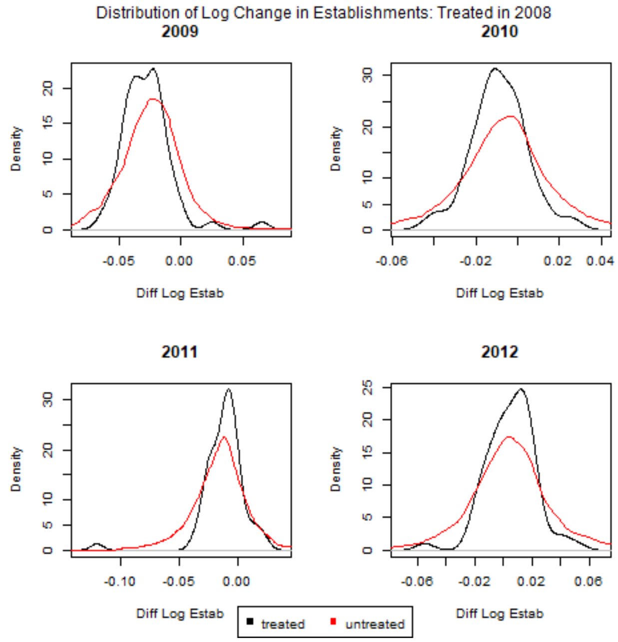
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
births2009	64	−0.028	0.020	−0.065	−0.040	−0.020	0.065
births2010	64	−0.009	0.013	−0.043	−0.016	−0.001	0.027
births2011	64	−0.012	0.019	−0.120	−0.021	−0.004	0.023
births2012	64	0.005	0.017	−0.055	−0.005	0.016	0.051
birthsNoCons2009	64	−0.019	0.019	−0.049	−0.032	−0.010	0.079
birthsNoCons2010	64	−0.003	0.012	−0.038	−0.010	0.003	0.033
birthsNoCons2011	64	−0.008	0.019	−0.120	−0.016	0.001	0.029
birthsNoCons2012	64	0.007	0.018	−0.059	−0.005	0.018	0.056
manufBirths09	64	−0.048	0.034	−0.131	−0.069	−0.031	0.047
manufBirths10	64	−0.023	0.036	−0.111	−0.046	0.000	0.058
manufBirths11	64	−0.024	0.052	−0.204	−0.048	0.000	0.095
manufBirths12	64	0.006	0.041	−0.112	−0.011	0.028	0.114
change.index.2008	64	−2.850	5.667	−33.460	−3.297	0.300	2.880
Subcomm..on.FI	64	0.125	0.333	0	0	0	1
Local.FIRE.donation	64	0.024	0.023	0.0001	0.007	0.033	0.087
Dem.	64	0.109	0.315	0	0	0	1
Treated.2009	64	0.000	0.000	0	0	0	0
Treated.2013	64	0.422	0.498	0	0	1	1
Treated.2008	64	1.000	0.000	1	1	1	1
Treated.2012	64	0.344	0.479	0	0	1	1
Unem.Rate	64	6.011	1.725	2.500	4.700	6.600	11.100
tier1ratio	64	1,199.963	870.403	39.000	509.500	1,751.750	3,500.000
troubledassetratio	64	264.818	336.583	1.000	1.000	524.500	1,277.000
ROA	64	1,768.472	848.323	30.000	1,198.000	2,300.500	4,117.000
cash.to.assets	64	2,230.687	1,180.018	150	1,432.4	3,085.2	5,100
loan.to.deposits	64	2,956.695	988.486	915.000	2,311.000	3,771.750	4,479.000
share_h	64	0.087	0.133	0.008	0.030	0.082	0.956
share_b	64	0.108	0.125	0.001	0.025	0.120	0.532
share_aa	64	0.027	0.025	0.003	0.008	0.035	0.097
treatFst.ex	64	0.000	0.000	0	0	0	0

Table 9: Treated 2009 Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
births2009	272	−0.025	0.021	−0.122	−0.037	−0.013	0.066
births2010	272	−0.007	0.019	−0.081	−0.017	0.002	0.063
births2011	272	−0.012	0.018	−0.115	−0.021	−0.003	0.062
births2012	272	0.006	0.023	−0.201	−0.003	0.017	0.067
birthsNoCons2009	272	−0.018	0.020	−0.107	−0.028	−0.008	0.087
birthsNoCons2010	272	−0.002	0.019	−0.086	−0.011	0.006	0.082
birthsNoCons2011	272	−0.009	0.019	−0.119	−0.019	0.001	0.081
birthsNoCons2012	272	0.007	0.026	−0.224	−0.003	0.020	0.069
manufBirths09	272	−0.058	0.057	−0.288	−0.082	−0.026	0.154
manufBirths10	272	−0.027	0.058	−0.288	−0.052	0.000	0.201
manufBirths11	272	−0.022	0.063	−0	−0.05	0.01	0
manufBirths12	272	0.015	0.083	−0	−0.02	0.04	1
change.index.2008	272	−2.001	5.502	−27.140	−3.465	0.968	19.140
Subcomm..on.FI	272	0.114	0.318	0	0	0	1
Local.FIRE.donation	272	0.021	0.018	0.000	0.007	0.031	0.098
Dem.	272	0.114	0.318	0	0	0	1
Treated.2009	272	1.000	0.000	1	1	1	1
Treated.2013	272	0.360	0.481	0	0	1	1
Treated.2008	272	0.000	0.000	0	0	0	0
Treated.2012	272	0.324	0.469	0	0	1	1
Unem.Rate	272	5.770	1.720	2.700	4.600	6.900	11.100
tier1ratio	272	1,275.935	699.704	126	703.1	1,736.1	3,509
troubledassetratio	272	434.753	485.503	1	1	700.8	2,010
ROA	272	2,001.501	1,121.218	26	1,102.8	2,776.9	5,142
cash.to.assets	272	2,294.843	1,051.362	40	1,558	3,006.8	5,036
loan.to.deposits	272	2,768.012	894.511	76	2,163.6	3,466.4	4,476
share_h	272	0.084	0.100	0.005	0.021	0.102	0.641
share_b	272	0.113	0.139	0.001	0.012	0.162	0.747
share_aa	272	0.024	0.042	0.0004	0.005	0.024	0.449
treatFst.ex	272	0.165	0.372	0	0	0	1

Table 10: Treated 2008 and 2009 Summary Statistics

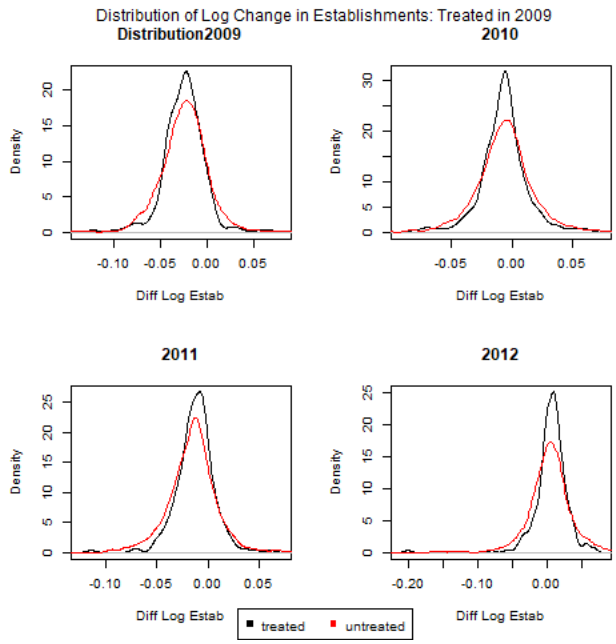
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
births2009	85	−0.020	0.014	−0.046	−0.029	−0.014	0.040
births2010	85	−0.005	0.011	−0.056	−0.012	−0.001	0.042
births2011	85	−0.010	0.012	−0.049	−0.018	−0.004	0.051
births2012	85	0.011	0.017	−0.025	0.001	0.019	0.117
birthsNoCons2009	85	−0.014	0.013	−0.036	−0.021	−0.009	0.058
birthsNoCons2010	85	−0.001	0.011	−0.019	−0.008	0.002	0.047
birthsNoCons2011	85	−0.007	0.013	−0.049	−0.015	0.0003	0.056
birthsNoCons2012	85	0.013	0.018	−0.022	0.003	0.021	0.123
manufBirths09	85	−0.055	0.031	−0.169	−0.072	−0.038	0.014
manufBirths10	85	−0.028	0.028	−0.105	−0.043	−0.014	0.065
manufBirths11	85	−0.022	0.027	−0.125	−0.037	−0.005	0.050
manufBirths12	85	0.008	0.033	−0.088	−0.007	0.026	0.107
change.index.2008	85	−4.444	6.859	−29.080	−5.390	−0.220	3.110
Subcomm..on.FI	85	0.212	0.411	0	0	0	1
Local.FIRE.donation	85	0.035	0.025	0.0003	0.019	0.044	0.131
Dem.	85	0.212	0.411	0	0	0	1
Treated.2009	85	1.000	0.000	1	1	1	1
Treated.2013	85	0.518	0.503	0	0	1	1
Treated.2008	85	1.000	0.000	1	1	1	1
Treated.2012	85	0.506	0.503	0	0	1	1
Unem.Rate	85	5.299	1.285	2.500	4.500	6.000	8.600
tier1ratio	85	1,629.000	861.530	169.500	1,040.500	2,087.143	3,524.000
troubledassetratio	85	390.485	406.876	1.000	69.000	537.000	1,967.000
ROA	85	1,750.622	1,062.274	88	987.4	2,443	5,105
cash.to.assets	85	2,531.211	1,103.249	226.000	1,768.000	3,365.000	4,958.000
loan.to.deposits	85	2,892.362	806.376	809.000	2,403.000	3,400.714	4,293.000
share_h	85	0.127	0.119	0.007	0.039	0.168	0.480
share_b	85	0.153	0.156	0.007	0.036	0.219	0.681
share_aa	85	0.050	0.061	0.005	0.016	0.052	0.339
treatFst.ex	85	0.435	0.499	0	0	1	1



7.3 Neighbor Characteristics

8 Technical Appendix

Decomposition of Matching Estimators



$$
\begin{aligned}
\hat{\tau}_t^{s_2{}^k, s_2{}^l}(\underline{s_2}^k) &= \frac{1}{N_{\underline{s_2}^k}, t} \sum_{D_{\underline{s_2}^k,t,i}=1} (Y_{it} - \hat{Y}_{it}(\underline{s_2}^l, G_{it}=0)) \\
&= \frac{1}{N_{\underline{s_2}^k}, t} \sum_{D_{\underline{s_2}^k,t,i}=1} (Y_{it} - \mu(\underline{x_{it}}, \underline{s_t}^k, g_{it}) + \mu(\underline{x_{it}}, \underline{s_t}^k, g_{it}) + \mu(\underline{x_{it}}, \underline{s_t}^l, g_{it}) - \mu(\underline{x_{it}}, \underline{s_t}^l, g_{it}) \\
&\quad - \frac{1}{M} \sum_{j \in \mathbb{J}_M(i,t,G_{it}=0, D_{\underline{s_2}^l,t,i}=1)} (Y_{jt} - \mu(\underline{x_{jt}}, \underline{s_t}^l, 0) + \mu(\underline{x_{jt}}, \underline{s_t}^l, 0) \\
&= \frac{1}{N_{\underline{s_2}^k}, t} \sum_{D_{\underline{s_2}^k,t,i}=1} (\epsilon_{it} + \mu(\underline{x_{it}}, \underline{s_t}^k, g_{it}) + \mu(\underline{x_{it}}, \underline{s_t}^l, g_{it}) - \mu(\underline{x_{it}}, \underline{s_t}^l, g_{it}) \\
&\quad - \frac{1}{M} \sum_{j \in \mathbb{J}_M(i,t,G_{it}=0, D_{\underline{s_2}^l,t,i}=1)} (\epsilon_{jt} + \mu(\underline{x_{jt}}, \underline{s_t}^l, 0) \\
&= \frac{1}{N_{\underline{s_2}^k}, t} \sum_{D_{\underline{s_2}^k,t,i}=1} \mu(\underline{x_{it}}, \underline{s_t}^k, g_{it}) - \mu(\underline{x_{it}}, \underline{s_t}^l, g_{it}) + \mu(\underline{x_{it}}, \underline{s_t}^k, g_{it}) - \mu(\underline{x_{it}}, \underline{s_k}^l, 0) \\
&\quad + \frac{1}{N_{\underline{s_2}^k}, t} \sum_{D_{\underline{s_2}^k,t,i}=1} \frac{1}{M} \sum_{j \in \mathbb{J}_M(i,t,G_{it}=0, D_{\underline{s_2}^l,t,i}=1)} (\mu(\underline{x_{it}}, \underline{s_t}^l, 0) - \mu(\underline{x_{jt}}, \underline{s_t}^l, 0) \\
&\quad + \frac{1}{N_{\underline{s_2}^k}, t} \sum_{i=1}^{N} (D_{\underline{s_2}^k,t,i}\epsilon_{it} - D_{\underline{s_2}^l,t,i} \frac{K(i,t,G_{it}=0, D_{\underline{s_2}^l,t,i}=1)}{M})\epsilon_{it}
\end{aligned}
$$

Table 11: Matching Estimator Algorithm for Total Effect

1. Delete all observations not belonging to $\underline{s_t}^k$ or $(\underline{s_t}^l, G_{i1} = 0, G_{i2} = 0)$
2. Define a weight $w_i^{\underline{s_t}^l} = 0$
3. Estimate a probit for

$$P(S_{i,1} = s_1^k, G_{i1} = g_1 \mid X_{0,i} = x_0, X_{0,-i} = x_{0,-i})$$
$$= P(S_{i,1} = s_1^k \mid G_{i1} = g_1, X_{0,i} = x_0, X_{0,-i} = x_{0,-i})P(G_{im1} = g_1)$$

   and calculate first-period probabilities.
4. Delete all obersvations in $\underline{s_t}^k$ with lower or higher first period probabilities than those in $\underline{s_t}^l$
5. For every observation left in $\underline{s_t}^k$, find the observation in $\underline{s_t}^l$ with the closet first-period probability
6. Estimate the second period propensity score

$$Pr(S_{i,2}^k = s_2^k, G_{i,2} = g_2 \mid X_{1,i} = x_{1,i},\ X_{0,i} = x_{0,i},\ S_{i1} = s_1, G_{i1} = g_1, X_{1,-i} = x_1',\ X_{0,-i} = x_{0,-i},\ S_{1,-i} = s_{-i})$$

   and calculate transition probabilities
7. Delete all obersvations in $\underline{s_t}^k$ with lower or higher second period transition probabilities than those in $\underline{s_t}^l$
8. For every observation left in $\underline{s_t}^k$, find the observation in $\underline{s_t}^l$ with the closet second-period probability

Table 12: Matching Estimator Algorithm for Main Effect

1. Delete all observations not belonging to $\underline{s_t}^k$ or $(\underline{s_t}^l$
2. Define a weight $w_i^{\underline{s_t}^l} = 0$
3. Estimate a probit for

$$
\begin{aligned}
&P(S_{i,1} = s_1^k, G_{i1} = g_1 \mid X_{0,i} = x_0, X_{0,-i} = x_{0,-i}) \\
&\qquad\qquad\qquad = P(S_{i,1} = s_1^k \mid G_{i1} = g_1, X_{0,i} = x_0, X_{0,-i} = x_{0,-i}) P(G_{im1} = g_1)
\end{aligned}
$$

   and calculate first-period probabilities.

4. Delete all obersvations in $\underline{s_t}^k$ with lower or higher first period probabilities than those in $\underline{s_t}^l$
5. For every observation left in $\underline{s_t}^k$, find the observation in $\underline{s_t}^l$ with the closet first-period probability, matching perfectly on link function outcome
6. Estimate the second period propensity score

$$
Pr(S_{i,2}^k = s_2^k, G_{i,2} = g_2 \mid X_{1,i} = x_{1,i},\ X_{0,i} = x_{0,i},\ S_{i1} = s_1, G_{i1} = g_1, X_{1,-i} = x_1',\ X_{0,-i} = x_{0,-i},\ S_{1,-i} = s_{-i})
$$

   and calculate transition probabilities

7. Delete all obersvations in $\underline{s_t}^k$ with lower or higher second period transition probabilities than those in $\underline{s_t}^l$
8. For every observation left in $\underline{s_t}^k$, find the observation in $\underline{s_t}^l$ with the closet second-period probability, matching perfectly on link function outcome

8.1 Matching Estimator Algorithms

Note, the relevant propensity scores changes based on which object in $\underline{s_t}^k$ is the match target. This requires estimating all possible sequences of transition probabilities based around the 16 possible period 2 potential outcomes.

Table 13: Matching Estimator Algorithm for Indirect Effect

1. Delete all observations not belonging to $\underline{g_t}^k$ or $(\underline{g_t}^l$
2. Define a weight $w_i^{\underline{g_t}^l} = 0$
3. Estimate a probit for

$$P(S_{i,1} = s_1^k, G_{i1} = g_1 \mid X_{0,i} = x_0, X_{0,-i} = x_{0,-i})$$
$$= P(S_{i,1} = s_1^k \mid G_{i1} = g_1, X_{0,i} = x_0, X_{0,-i} = x_{0,-i})P(G_{im1} = g_1)$$

   and calculate first-period probabilities.

4. Delete all obersvations in $\underline{g_t}^k$ with lower or higher first period probabilities than those in $\underline{g_t}^l$
5. For every observation left in $\underline{g_t}^k$, find the observation in $\underline{g_t}^l$ with the closet first-period probability, matching perfectly on link function outcome
6. Estimate the second period propensity score

$$Pr(S_{i,2}^k = s_2^k, G_{i,2} = g_2 \mid X_{1,i} = x_{1,i},\ X_{0,i} = x_{0,i},\ S_{i1} = s_1, G_{i1} = g_1, X_{1,-i} = x_1',\ X_{0,-i} = x_{0,-i},\ S_{1,-i} = s_{-i})$$

   and calculate transition probabilities

7. Delete all obersvations in $\underline{g_t}^k$ with lower or higher second period transition probabilities than those in $\underline{g_t}^l$
8. For every observation left in $\underline{g_t}^k$, find the observation in $\underline{g_t}^l$ with the closet second-period probability, matching perfect