Studies of Social and Commercial Benefits of Postal Services:

ECONOMIC EFFECTS OF POST OFFICES

Final

August 2011

Submitted to:

Postal Regulatory Commission
901 New York Avenue, NW
Suite 200
Washington, DC 20268-001

Contract No. 109909-10-Q-0017
UI No. 08557-000-00

Authors:

Christopher Hayes
Christopher Narducci
Nancy Pindus

Advisors:

George Galster
Doug Wissoker

URBAN INSTITUTE

2100 M Street, NW ● Washington, DC 20037

TABLE OF CONTENTS

Introduction

This study uses a difference-in-differences approach to assess the economic impact of a post office closure on a community. We caution readers that this is an initial step toward measuring the impacts of post offices in communities. We use United States Postal Service (USPS) administrative data provided by the Postal Regulatory Commission[1] on locations of publicly accessible offices that are open and offices that closed between 2002 and 2005. Using Census data[2] on businesses, which include the total number of business establishments and employees, we matched ZIP codes that experienced a closure with geographically and economically comparable ZIP codes that have an open post office. Using several variants of our model, we found more variation in estimated impacts than we had hoped, but our results suggest a small, sometimes significant, negative impact on employment in the ZIP codes with post office closures. The weakness of the models is reflected in the confidence intervals around the results. This suggests that future research would benefit from larger samples that include additional control variables in order to test this finding more rigorously.

This report summarizes the research design, data, and methods used to arrive at the results reported. We then discuss the findings, significance, and limitations of the analysis. Appendix 1 provides the regression results. Appendix 2 briefly describes an alternative model that we tried, and Appendix 3 details data processing methods for preparing the data and drawing the sample.

Research Design

Determining economic impacts on communities poses difficult evaluation challenges. Communities are not static, so during any time period under investigation there are many other contextual factors beyond the existence or closure of a post office that may affect outcomes. These include changes in labor market conditions (such as business openings or closings), spatial factors (such as the building of a new road or exit ramp), and other investments—public or private. Not only must data regarding outcomes be collected but, also, some estimate of what outcomes would have occurred absent the intervention (in this case the post office or post office closure) must be made. Experimental designs do this by dividing subjects at random into two groups—one participating in a program or intervention and one not. Such methods are infeasible with respect to the study of post office impacts, where neither the location of post offices nor the closure can occur randomly.

[1] File constructed by USPS. See Data section for details.

[2] This excludes post offices and postal workers, as well as most government employees. http://www.census.gov/econ/cbp/overview.htm.

Lacking the conditions for a classical random assignment experimental design, we used a design that enabled us to attribute various outcomes to the presence of a post office and statistically estimate the likelihood of these associations. We compared communities or neighborhoods with a post office to similar communities or neighborhoods that had a post office that was closed between 2002 and 2005. Our hypothesis is that the closing of a post office has negative effects on the economic indicators of the surrounding community. In our design, it is the difference between the pre-post change of a community with a post office closure and similar communities without such a closure that provide our evidence of impact. The comparisons must be carefully constructed and interpreted. For example, in considering the impact of investments aimed at revitalizing low-income neighborhoods, Galster, Tatian, and Accordino[3] note that comparisons between declining neighborhoods may not show a reversal of the decline due to an economic development intervention, but that one should not overlook a slowing in economic decline as evidence of positive impact. Furthermore, citywide factors may affect economic outcomes in all neighborhoods.

We applied the well-accepted "difference-in-differences" model, which takes advantage of the panel data available. For each ZIP code with a closed post office, we selected two matched ZIP codes that did not lose their post offices, either from the same metro area, or, in the case of non-urban ZIP codes, rural ZIP codes in the same state. After selecting the matching ZIP codes, we then calculated pre-post changes in local employment and number of business establishments for areas with and without the loss of a post office and then compared the pre-post changes in areas with closed post offices to those of matched communities.

The logic of the difference-in-differences model is as follows. All the members of the triad of ZIP codes being compared are assumed to be subject to roughly the same macroeconomic forces that would similarly affect their trends in the outcome indicators over the analysis period. To the extent that members of these triads are not perfectly matched during the pre-closure period, there may be initial differences among them. However, to the degree that a post office closure affects one of the ZIP codes in the triad, there will be a *difference* in these initial differences that appears during the post-closure period. It is this difference-in-differences that represents our best estimate of the causal impact of the post office closure or, equivalently, the value of an operating post office to a ZIP code.

In this approach we compared closure ZIP codes only to those with similar geographic and economic qualities. We first limited potential matches to those ZIP codes within the appropriate geography—the same metropolitan area for urban ZIP codes or non-metro ZIP codes located in the same state for rural ZIP codes. We next restricted potential matches to ZIP

[3] George Galster, Peter Tatian, and John Accordino, "Targeting Investments for Neighborhood Revitalization," *Journal of the American Planning Association* 72, No. 4: 457–74.

codes with similar overall levels of number employed or number of establishments in 2000. To do this, we ranked ZIP codes by the number of establishments or employees, and selected only ZIP codes within the same ranking category as possible matches. ZIP codes were typically ranked into six categories, though this parameter was relaxed for several ZIP codes for reasons discussed below.

The final two matched ZIP codes were then chosen to be, to the extent possible, on a similar pre-closure trajectory in terms of our outcome indicators. That is, matches were chosen so that the rate of change of total employees or total establishments during the pre-closure period was comparable to the rate of change in our closure ZIP code. We attempted to select the two ZIP codes with the closest absolute rate, without regard to whether it was above or below the closure ZIP code rate. Given the limitations of each selection pool, this provided the best and most efficient means of matching ZIP codes. In sum, matched ZIP codes were from comparable geographies, had the closest base number of the outcome variables (establishment or employee) in 2000, and had the closest rate of change of this outcome between 2000 and the year of closure.

Following each selection of two matched ZIP codes for a closure ZIP code, we eliminated these ZIP codes from the selection pool of potential matches for other closure ZIP codes.[4] We also relaxed the ranking criteria discussed above for several closure ZIP codes in order to ensure matches. While it would not necessarily be fatal to the model to use a comparison ZIP code more than once, the relatively high rate of overlap (approximately 20 percent of our sample in preliminary trials) raised concern that the results could be biased. Replacement ZIP codes may not have matched the closure ZIP codes as closely as original selections, but in all cases a sufficient sample was available to provide a reasonably close match according to our parameters.

Selecting matching ZIP codes for each closure ZIP code adds to the face validity of the model and offers some controls for non-included factors. We stress, however, that the model's accuracy does not rely on perfect matches; it is the difference in the differences between the matched ZIP codes and the ZIP code with the closure observed before and after the closure that matters. We believe that we have established matches for ZIP codes with post office closures that are adequate for our model.

[4] We would have preferred to select best-match ZIP codes with trajectories slightly above and slightly below our closure ZIP codes. However, we found that some potential comparison ZIP codes in this scenario would have matched on more than one closure ZIP code, and so we adjusted the strategy to eliminate this potential overlap.

Data

We used the following data sources in the impact model:

- Decennial Census 2000 (demographic, economic, housing, and social characteristics)
- Information on business establishments and employment from the Census Bureau's ZIP Business Patterns database, 2000 to 2008
- Addresses for all open USPS facilities, 2009[5]
- Lists of all closed or suspended post office facilities, provided by the Postal Regulatory Commission (PRC), 2002 to 2005

The model uses ZIP codes as the geographic unit of analysis. Our list of closed postal facilities provides no street address, and our outcome indicators are most readily available at the ZIP code level. Since we are primarily concerned with neighborhood effects, ZIP code-level analysis should be adequate, though future studies might better assess impact through analysis of standardized longitudinal data at smaller geographic levels if such were to become available.

We selected indicators from ZIP Business Patterns for our outcome measures as the best gauges of business activity available yearly over our performance period at the ZIP code level. Reliable home price indicators, which would have been useful as a measure of neighborhood well-being, were not available on a yearly basis at this geographic scale across the nation.

We drew our sample of ZIP codes with closed facilities from the universe of closed post offices provided by the PRC.[6] In order to provide sufficiently long pre-closure and post-closure periods, we determined that we would restrict the list to facilities closed between 2002 and 2005. While the list of USPS facilities closed over our performance period included 245 post offices, we further reduced the total by eliminating ZIP codes which also had a currently open postal facility and those for which we lacked outcome data. We also selected only postal branches and stations, because of some evidence that other categories could include facilities not open to the public. Exploration of the category labeled "Post Office" revealed a number of facilities that appeared to have no public access (such as distribution centers). This is undesirable for our model, since our hypothesis is that the presence of the postal facility draws foot traffic to the area, which is then more likely to patronize local businesses. If a post office is

[5] Provided to the Urban Institute by Doug Carlson. The files were generated by the USPS for Mr. Carlson after a lawsuit established his right to request the data under the Freedom of Information Act.

[6] The data file on closed postal facilities is from testimony provided by the USPS to the PRC on August, 28, 2009, Library reference: USPS-LR-N2009-1/10 (Docket No. N2009-1).

closed to the public, it would not have this impact on surrounding businesses under these assumptions.

Our list of currently open facilities comes from late 2008. We attempted to ensure comparability with the closed offices by restricting the comparison group to offices that have lobby hours. Some facilities on this list may have been opened during or after our performance period; if so, these random errors will weaken the precision of the results. For instance, ZIP codes where a post office closed would effectively be compared to ZIP codes where a post office did not exist until after the period of interest. After examination of this issue, we concluded that the impact would be minimal: Very few facilities were opened in the previous decade; we would be selecting a very small sample from the total number of facilities (152 out of several thousand); and historical information on facility openings is irregular and difficult to extract.

In a few ZIP codes, values for the outcome variable reported in the business pattern data were extreme outliers. Before selecting our sample, ZIP codes containing values more than 6 standard deviations above the mean were excluded. The pool of closure ZIP codes was 86 after this process of eliminating unlikely values, consisting of roughly twice as many rural ZIP codes as urban.

However, this procedure left several ZIP codes with extreme values in the pool, which, while potentially accurate, were still well beyond the normal range of values. We therefore created a second pool of ZIP codes, selecting a lower threshold for inclusion. The final threshold for inclusion for number of employees in 2000 was 1,000 or fewer. We also restricted the pool of ZIP codes, both closure and matching, by eliminating any ZIP code where the year 2000 value was 0. Because our models are attempting to determine the effect of closure on levels of employment and business activity, ZIP codes that began our analysis period without measurable activity were considered to be unsuitable for inclusion. These restrictions reduced our list of closure ZIP codes to 69, a total of 51 in rural areas and 18 in urban areas. Both the larger and smaller sample pool were tested in the impact models—the more narrowly restricted sample provides a sounder pool with which to measure outcomes, but at the cost of sample size.

After restricting our sample of closure ZIP codes, analysis of selected indicators reveals stark differences between ZIP codes where closures happened and where they did not. Table 1 shows the differences between our selected closure ZIP codes and the pool of all ZIP codes with facilities open to the public for our outcome measures and selected socioeconomic indicators. The outcome indicators in particular demonstrate the substantially different character of the closure ZIP codes. There is less variation by socioeconomic characteristics, although the closure ZIP codes tend to be substantially poorer. While the steps described above to reduce the pool of the closure ZIP codes resulted in lower average numbers of employees and establishments, no difference is apparent in socioeconomic characteristics. Because both

minorities and the elderly represent potentially vulnerable populations that could be disproportionately affected by the loss of an easily accessible postal facility, it is encouraging that, in general, closures do not appear to be occurring in ZIP codes where these groups are concentrated.

Table 1. Characteristics of ZIP Codes
ZIP Codes with open facilities vs. ZIPs where a facility was eventually closed

	All ZIPs with public POs	Closure ZIPs	Restricted set of closure ZIPs
Number of	35,224	86	69
Mean establishments (2001)	201	21	6
Mean employees (2001)	3,014	553	36
Unemployment rate (2000)	6%	10%	10%
Poverty rate (2000)	12%	26%	26%
Percent minority (2000)	30%	20%	21%
Percent elderly (2000)	13%	12%	12%

Source: 2000 Decennial Census, U.S. Bureau of the Census, and ZIP Business Patterns, U.S. Bureau of the Census.

Matched ZIP codes were selected for the pool based on pre-closure outcome data, so averages for our pooled matches on those indicators are close to the closure ZIP code average, as we would expect. Some differences remain between the groups, however, on the socioeconomic indicators. The impact of these differences on the model would be worth exploring in future research.

Methods

Difference-in-Differences Model

We tested three different techniques for running difference-in-differences models in order to assess the consistency of our results. First, we used a Generalized Least Squares (GLS) regression procedure with random effects by triad of matched ZIP codes, which reweights the data to provide efficient estimates, assuming that the econometric model is specified correctly. Then we ran a model with the same set of predictors, using a standard Ordinary Least Squares (OLS) regression. This strategy makes weaker assumptions than the

random effects model, although it lacks any efficiency gain resulting from its longitudinal features. We used robust clustered standard errors in both approaches, which allow correlation among the errors within the triads selected. Finally, we tested a fixed effects regression model. This controls for factors not included in the model that may affect total employment and that differ across ZIP codes, but are constant over time for each ZIP code. In this last approach, the difference-in-differences relies on the difference in the change over time in employment between ZIP codes with a closing and the matched comparison areas.

For each approach, we modeled the number of employees over time in each set of triads.[7] We ran the models using both the level and log of the number of employees as the dependent variable. The purpose of modeling the log is to reduce the effect of relatively large values of employment in the calculations. In contrast to the level (linear) model, the coefficients in a model with a logged value produce a coefficient interpreted as the proportional rather than absolute change in numbers of employees associated with a one unit change in the given independent variable. The only difference in the equation specification for these runs is the logged dependent variable.

For both the GLS and OLS methods, we tested several variations of the same model. The fixed effects model included a separate intercept for each ZIP code, but not the set of control variables measured in the year 2000. The control variables cannot be included in the same model as the fixed effects since all of the variation in the control variables is captured by the ZIP code intercepts. The ZIP code intercepts controlled more completely for the differences in the average employment within each site. The inclusion of ZIP code dummies produced results with a higher r-squared than the other two methods, which should not be interpreted as a more powerful explanation of change over time (see the Findings section, below).

1. *Base GLS model specification, number of employees**

$$\boldsymbol{Totalemp_{it} = \beta_0 + \beta_1 closed_i + \beta_2 post_{it} + \beta_3 (post \cdot closed)_{it} + \delta_i + \varepsilon_{it}}$$

2. *Base OLS model specification, number of employees**

$$\boldsymbol{Totalemp_{it} = \beta_0 + \beta_1 closed_i + \beta_2 post_{it} + \beta_3 (post \cdot closed)_{it} + \varepsilon_{it}}$$

** In the GLS model,* δ *represents the random effect for the triad of which zip code i is a member. In both models, the subscripts indicate ZIP code* i *for time period t.*

[7] As a secondary test of our results, we also selected ZIP code matches based on the number of establishments and ran the same model for these triads as well. Theoretically, the number of employees would be a better outcome variable for the purposes of this study. This number is more sensitive to slight changes in economic conditions for the obvious reason that a business is more likely to reduce its staff before closing entirely. While we tested both outcomes, the number of employees is used in our primary test.

Our simplest model regressed the number of employees in a ZIP code on several one-zero indicator variables (equations 1 and 2, above). These included whether a ZIP code experienced a closure ("*closed"*), whether the outcome variable for members of the triad was observed in a year after the closure occurred ("*post"*), and an interacted variable of closure in the ZIP code and post-closure year ("*post-closed"*). This latter variable provided the estimate of the difference-in-differences effect—measuring the average difference between the two ZIP codes with open post offices and that with a closed post office in the years following the closure. Timing for the post-closure term in the above equation (*"post"*) for ZIP codes with an open post office is based upon the post office closure in the ZIP code with which they are matched. The fixed effect specification is similar to this base model other than that the inclusion of a dummy variable for each ZIP code replaces the variable *closed*.

Pre-closure characteristics may also have explanatory power that could strengthen the model. Therefore, we modified the model by sequentially adding new independent variables to control for conditions related to socioeconomic characteristics in the ZIP code prior to the post office closure (equations 3 and 4). These included median income, unemployment, percent foreign born, housing unit density, and the poverty rate. We also incorporated time indicator variables for years between 2000 and 2008 to control for changes across years ("d00"–"d08"). Nevertheless, it is possible that unobserved, time-varying ZIP code-level conditions affected the results of the basic model, although we selected ZIP codes with similar pre-closure trajectories to reduce the chances of this problem.

3. GLS Model specification including controls, number of employees

$$\boldsymbol{Totalemp_{it} = \beta_0 + \beta_1 closed_i + \beta_2 post_{it} + \beta_3 (post \cdot closed)_{it} + \beta_4 med\_income_i + \beta_5 unemply_i + \beta_6 pct\_min_i + \beta_7 hsg\_dens_i + \beta_8 pct\_forgn_i + \beta_9 poverty_i + \beta_{10} \mathbf{d00}_t + \beta_{11} \mathbf{d01}_t + \beta_{12} \mathbf{d02}_t + \beta_{13} \mathbf{d03}_t + \beta_{14} \mathbf{d04}_t + \beta_{15} \mathbf{d05}_t + \beta_{16} \mathbf{d06}_t + \beta_{17} \mathbf{d07}_t + \beta_{18} \mathbf{d08}_t + \delta_i + \varepsilon_{it}}$$

4. OLS Model specification including controls, number of employees

$$\boldsymbol{Totalemp_{it} = \beta_0 + \beta_1 closed_i + \beta_2 post_{it} + \beta_3 (post \cdot closed)_{it} + \beta_4 med\_income_i + \beta_5 unemply_i + \beta_6 pct\_min_i + \beta_7 hsg\_dens_i + \beta_8 pct\_forgn_i + \beta_9 poverty_i + \beta_{10} \mathbf{d00}_t + \beta_{11} \mathbf{d01}_t + \beta_{12} \mathbf{d02}_t + \beta_{13} \mathbf{d03}_t + \beta_{14} \mathbf{d04}_t + \beta_{15} \mathbf{d05}_t + \beta_{16} \mathbf{d06}_t + \beta_{17} \mathbf{d07}_t + \beta_{18} \mathbf{d08}_t + \varepsilon_{it}}$$

Findings, Significance, and Limitations

We first discuss the results of our models for impacts on the number of employees. All models produced a similar negative magnitude of impact from a post office closure of roughly six jobs lost in the ZIP code, with modest variation across the models in standard errors and statistical significance. The alternative models led to similar point estimates, with significance levels slightly above and slightly below traditional minimum standards of significance. As we added control variables to our GLS model, we did not see much added strength of the model or of the significance of our difference-in-differences variable's coefficient. When running a standard OLS, the results were similar. Even so, neither estimate of impact was statistically significant at conventional levels. When running a fixed effects regression, however, the model's significance was greatly improved. The impact variable coefficient was consistent with the other methods, at approximately six jobs lost, and statistically significant at the 0.05 level of confidence. However, the increased significance may be caused by the lack of adjustment of the standard error due to clustering in this specification, or simply by the shrinking of the standard error of the regression resulting from inclusion of the dummies for ZIP codes (Appendix 1).

Our results using the log of number of employees display similar patterns, though none produce findings that are statistically significant at near conventional levels. In neither the GLS nor OLS models does the addition of control variables affect the estimated difference-in-differences. This is likely because the control variables are measured only in the initial period. The coefficient on our impact variable is negative and consistent across all specifications, however. The fixed effects method produces the finding for which we are most certain of a negative effect, but again, this is likely due to the lack of adjustment of the standard error due to clustering in this specification. While the impact variable coefficient is consistent with the other models in both magnitude and direction, the confidence interval around the estimate is quite large, and it is not significant at conventional levels.[8]

Overall, use of the constrained sample led to a relative consistency in estimating the impact across the model types and specifications tested, and gave us supportable indications of a weak but negative impact on the ZIP codes in our sample. We would advise caution in drawing conclusions about impact in areas with greater numbers of employed individuals than in this constrained sample.[9] Furthermore, these results are an average effect estimated over a

[8] The estimated coefficient is -0.35, or a loss of 30 percent of employees, but the true effect may be anywhere between a decrease of 64 percent and an increase of 33 percent.

[9] We originally ran these models including ZIP codes with an initial value of 0, and with the higher tolerance for extreme values (see Data and Methods sections), and these produced estimates that depended heavily upon whether GLS or OLS was used for the estimation, with high variation in confidence levels. We believe this variation resulted from the presence of extreme outliers, since excluding the outliers or using log-dependent variables both led to more consistent findings across specifications. As a result, we decided to constrain our sample as described above.

wide range of ZIP code sizes. For some ZIP codes, a loss of six jobs would represent little change, while for others (i.e., those with five employees), it would be devastating.

Implications/Suggestions for Further Research

The results of this study do not provide conclusive evidence of economic impact, but they do suggest that future research on the relationship between post offices and business activity is warranted. Our model utilized ZIP code-level indicators due to research constraints: we were unable to locate national, longitudinal data at smaller geographies. Because ZIP codes vary in size and are, on average, larger than the area where the effects of business agglomeration would be felt, it is likely that indicators at smaller neighborhood levels proximate to post offices that are closed are better suited for this type of study. The size variation of ZIP codes may be a driving force behind the large standard errors. Further, if economic effects do result from post office closures, these effects are likely to be stronger at the neighborhood or block level proximate to post offices that are closed.

Future studies should also attempt to increase the sample size of closed post offices. We limited our test to ZIP codes that experienced a closure during the period from 2002 to 2005, but a sample that includes closures from additional years would produce more reliable results. Researchers could also draw from a wider variety of facility types, if resources were available to verify that the facility was open to the public pre-closure (this information was not included in currently available data).

Beyond improving the accuracy of a full sample test, expanding the sample would allow researchers to run reliable tests on subgroups such as urban, rural, low-income, and high-minority communities. It is likely that including all of these subtypes in our regression of ZIP codes resulted in large standard errors, yet our sample was not sufficient to stratify outcomes by these typologies.

APPENDIX 1:
Regression Results

Results for the tests of impact on levels of employees ("Number of employees," above), for all three types of models and both base and full specifications indicate a consistent negative effect due to post office closure in the affected ZIP code.

- For both GLS specifications the coefficient is -6.3, but is not significant and has a confidence interval that indicates the actual value is somewhere between a decrease of 14.6 employees and an increase of 1.9 employees.
- The coefficient for the OLS base specification is -9.8, which is significant at the 0.05 level, but once again the confidence interval is large, with the actual value likely falling between -19.5 and -0.2.
- The coefficient for the full OLS specification is not significant.
- The coefficient for the fixed effects approach is significant, and has a smaller confidence interval than the other two approaches, but as explained in the main body of the report, this may be due to the lack of adjustment of the standard error due to clustering in this specification, or the inclusion of dummies for ZIP codes.

Other terms in the model indicate changes in employment over time and differences between closure ZIP codes and open ZIP codes overall. The post-closure period across all ZIP codes, closure and non-closure, is associated with a small increase in employment, significant at the 0.05 level in the OLS specifications. The actual value for this impact most likely falls within a wide range, from an increase between 1.9 and 23.4 employees according to the base OLS model, to an increase between 8.0 and 74.6 employees according to the full OLS specification. Also, even though non-closure ZIP codes were selected for the similarity to closure ZIP codes in employment before the closure, estimates for closure ZIP codes indicate significantly lower levels of employment from 2000 to 2008 compared with the matched ZIP codes and independent of losses after closure.

For the models estimating the impact of closure on the logged number of employees, no estimates were significant. Coefficients were negative, but standard errors for all specifications were high. The confidence intervals for actual values include a potential increase in the number of employees.

CONSTRAINED SAMPLE
(ZIP Codes with zero employees or more than 1,000 employees in 2000 excluded)

	Base specification			Full specification	
	GLS	OLS	FE	GLS	OLS
Number of employees					
Closure ZIP, post-closure	-6.35	-9.82 **	-6.32 **	-6.33	-7.86 *
	(4.213)	(4.839)	(2.904)	(4.224)	(4.604)
All ZIPs, post-closure	2.31	12.66 **	1.37	1.88	41.28 **
	(2.580)	(5.366)	(3.194)	(3.186)	(16.694)
Closure ZIP	-13.02 ***	-11.28 ***		-16.64 ***	-15.91 ***
	(4.535)	(4.011)		(5.922)	(6.033)
Median HH income, 2000				0.003	0.003
				(0.0027)	(0.0027)
Unemployment rate, 2000				0.480	0.417
				(0.514)	(0.504)
Pct minority, 2000				-0.749 **	-0.797 **
				(0.326)	(0.343)
Res. housing density, 2000				0.147 **	0.145 **
				(0.068)	(0.0669)
Pct foreign born, 2000				1.87	1.76
				(1.303)	(1.255)
Poverty rate, 2000				1.13	1.20
				(1.254)	(1.249)
_cons	52.04 ***	46.83 ***	51.47 ***	-68.852	-70.954
	(11.866)	(10.348)	(2.004)	(97.133)	(96.948)
Logged number of employees					
Closure ZIP, post-closure	-0.3633	-0.4635	-0.3549	-0.3560	-0.3685
	(0.330)	(0.360)	(0.221)	(0.331)	(0.366)
All ZIPs, post-closure	-0.3103 *	-0.2656	0.1027	0.1005	0.0810
	(0.184)	(0.225)	(0.243)	(0.301)	(0.596)
Closure ZIP	-0.9649 ***	-0.9144 ***		-1.1699 ***	-1.1635 ***
	(0.257)	(0.265)		(0.30)	(0.306)
Median HH income, 2000				0.0000	0.0000
				(0.00003)	(0.00003)
Unemployment rate, 2000				0.0018	0.0019
				(0.0354)	(0.0354)
Pct minority, 2000				-0.0651 ***	-0.6509 ***
				(0.0212)	(0.0210)
Res. housing density, 2000				0.0029 **	0.0029 **
				(0.0012)	(0.0012)
Pct foreign born, 2000				0.1055 **	0.1056 **
				(0.044)	(.0441)
Poverty rate, 2000				-0.0112	-0.0112
				(0.027)	(0.0274)
_cons	2.4149 ***	2.3923 ***	2.7230 ***	1.9264	1.9250
	(0.242)	(0.228)	(0.152)	(1.434)	(1.435)

*** p<0.01, ** p<0.05, * p<0.1

Appendix 2:
Synthetic Control Model

Earlier in this project, we experimented with a new modeling technique, known as the synthetic control method, which substituted for an actual matched ZIP code an artificial matched case.[10] This is accomplished by averaging indicators across many cases instead of using the hand-selected matches described above. This method has the potential to overcome difficulties caused when good matches in the non-treatment pool cannot be found.

In our application of the model, use of a complex weighting algorithm was intended to select cases that could contribute to an ideal matching case (based on pre-closure data), whose post-closure trajectory could then be compared to the actual closure ZIP code. While this approach seemed to hold some promise, we found that it was unable to construct synthetic ZIP codes for the comparison that met goodness-of-fit criteria. While some tests produced seemingly sound synthetic cases, for most ZIP codes the synthetic case was not a good match. Closer examination of the few good fits, however, led us to be skeptical of those results as well. We decided that the preferred approach was the difference-in-differences model described in this report.

[10] Alberto Abadie, Alexis Diamond, and Jens Hainmueller, "Synthetic control methods for comparative case studies: Estimating the effect of California's Tobacco Control Program" (January 2007), http://www.people.fas.harvard.edu/~jhainm/Paper/ADH2007.pdf.

APPENDIX 3:
Data Processing Methods

Closure ZIP Codes

We drew our sample of ZIP codes with closed facilities from the universe of closed post offices provided by the PRC. The data file on closed postal facilities is from testimony provided by the USPS to the PRC on August, 28, 2009, Library reference: USPS-LR-N2009-1/10 (Docket No. N2009-1). We limited this data set to those closure ZIP codes where the "Actual_Date" close date is between 2002-2005 and where "DIS_POTYPE" is either A (Post office), C (classified station), or D (classified branch). We excluded retail contract offices from our list, as well as those offices with no lobby hours listed. In cases where ZIP codes experience more than one post office closure, we eliminated all but one ZIP code observation.

Comparison ZIP Codes

Our pool of ZIP codes with open facilities from which matched comparisons were selected came from a 2008 set provided by Doug Carlson (unconnected to this project or the Urban Institute), who obtained the records through a Freedom of Information Act request. Information on facilities available from the file included name, address, type and subtype, and window and lobby hours by day of week. We limited this to those open ZIP codes where "Facility_Subtype" is "MAIN_PO," "STATION," "BRANCH," "CPU_B," "CPU_C," "CPU_S," or "FIN_S," and the facility has lobby hours. These include all offices that are open to the public.

Independent and Dependent Variables

We drew ZIP code level variables to incorporate into our regression model from the 2000 Decennial Census and ZIP Business Patterns 2000 through 2008.

Our 2000 Census variables include:

- Total Population
- Population Density
- Population Minority Alone
- Unemployment Rate
- Median Household Income
- Number of Housing Units
- Poverty Rate
- Population Aged 65 Over
- Land Area (Square Miles)
- Population Foreign Born
- ZIP Code

We created the following variables using this Census data.

- Percent Minority
- Percent Foreign
- Percent Elderly

Housing Unit Density

Our 2000–2008 ZBP variables include:

ZIP Code
State
Total Establishments
Total Employees
Year

Universe of Potential Comparison ZIP Codes for Further Processing

We merged the closure ZIP codes, open ZIP codes, and accompanying ZIP Business Patterns (ZBP) and Census data. Using 2000 Census data that link ZIP codes with Core-Based Statistical Area (CBSA) codes, we defined metropolitan and non-metropolitan ZIP codes. The ZIP codes that are not in a metropolitan area have a CBSA code of 999999 and are defined as rural.

From this data set, we eliminated all ZIP codes that were not on our closure or open lists and all ZIPs that lacked ZBP/Census data. We also eliminated all closure or open ZIP codes where the total employment in 2000 was equal to 0.

Creating Comparison Triads

For each closed ZIP code:

1. We selected all open ZIP codes in the same Metropolitan Statistical Area (MSA) based on the CBSA code. If a closure ZIP code was not in an MSA, we selected all non-MSA ZIP codes from the same state.
2. From this pool, we ranked the closure ZIP codes and open ZIP codes among equal categories according to the total number of employees in the year of closure of the respective closure ZIP code. Then we constrained the selection pool by selecting only ZIP codes that were in the closure ZIP code's same ranking category. For 79 closure ZIP codes, the comparison pool was ranked into 6 categories before selection. For 7 closure ZIP codes, the comparison pool was ranked into 4 categories. For the final single closure ZIP code, selection based on ranking prevented it from matching with two comparison ZIP codes, so this step was skipped entirely.
3. For the remaining ZIP codes in the comparison pool, we calculated the rate of change of total employees between 2000 and the date of closure. Because ZPB surveys are collected in early March, if a post office is closed in January or February, the rate for closure and comparison pool ZIP codes was calculated using the previous year's employment data. We selected the two ZIP codes with the nearest absolute rate change to create our final triad.
4. We assigned the closed ZIP code and the two matched ZIP codes a unique triad ID.
5. After each triad was created, we eliminated the matched ZIP codes from the universe of potential comparison ZIP codes before running the procedure for the next closure ZIP code. In other words, each progressive run of the matching procedure eliminated two open ZIP codes from the initial pool of the next closure ZIP code to be matched.
6. We merged all triads into a single data set that included ZBP, Census, and created indicator variables.

7. We flagged all triads for which the closure ZIP code had greater than 839 employees in 2000. Analysis of the distribution of total employees determined that ZIP codes with more than 839 employees were extreme outliers. These were selected out of our final analysis.

The final data set used for this study included 69 triads consisting of a closure ZIP code and two unique matched comparison ZIP codes with open facilities. Following is a table of relevant variables included in this data set and their definitions.

Variable Name	Description
id	Triad ID
ZIP	ZIP code
TotalEmp	Total employment (ZPB)
TotalEstab	Total establishments (ZPB)
ne0	Total employment in 2000 not equal to 0 (0/1)
MedianHshldIncome_2000	Median income 2000 (Census)
UnemploymentRate_2000	Unemployment rate 2000 (Census)
PovertyRate_2000	Poverty rate 2000 (Census)
PopDensity_2000	Population density 2000 (Census)
year	Year of observation
Minority	Percent not non-Hispanic white 2000 (Census)
Foreign	Percent foreign-born 2000 (Census)
Elderly	Percent over age 65 2000 (Census)
HsgDens	Housing density
actual_date	Actual date of closure
STATE	State
closed	Closure ZIP (0/1)
close_year	Year of closure
close_month	Month of closure
CBSA	CBSA code
metro	ZIP lies in an MSA (0/1)
totalemp_close	Total employees in year of closure
pct_chg_totalemp	Percent change of total employees 2000 to year of closure
post_closeyr	Observation after year of closure (0/1)
postXclose	Observation of closure ZIP after year of closure (0/1)
d00	Observation from 2000 (0/1)
d01	Observation from 2001 (0/1)
d02	Observation from 2002 (0/1)
d03	Observation from 2003 (0/1)
d04	Observation from 2004 (0/1)
d05	Observation from 2005 (0/1)
d06	Observation from 2006 (0/1)
d07	Observation from 2007 (0/1)
d08	Observation from 2008 (0/1)
emp839	Total employees in year of closure greater than 839 (0/1)