Canadian Labour Economics Forum

WORKING PAPER SERIES

The Effects of Field Office Closures on Social Assistance Take-Up and Targeting

Jeffrey Hicks (University of Toronto)

CLEF WP #87

The Effects of Field Office Closures on Social Assistance Take-up and Targeting

Jeffrey Hicks*

August 9, 2025

Abstract

How does in-person support affect access to safety net programs? I study this question by examining how the closure of field offices has changed welfare and disability assistance caseloads. Using rich administrative data and a staggered difference-in-differences design, I estimate that closures, on average, reduced local caseloads by 11.5% for welfare and a statistically insignificant 1.6% for disability assistance. Declines in welfare caseloads (i) occurred across demographic, health, education, and eligibility groups, (ii) were somewhat larger among young and healthier individuals, and (iii) were suggestively larger among persons less familiar with the programs. On the whole, I find limited change in the relative targeting of benefits.

JEL Codes: H31, H83, H75, H53, I32, I38, J18.
Keywords: welfare, disability assistance, screening, take-up.

Disclaimer: All inferences, opinions, and conclusions drawn in this paper are those of the authors and do not reflect the opinions or policies of the Data Innovation Program or the Province of British Columbia. The Data Innovation Program data sets used in this study are listed in Section 6. You can find further information regarding these data sets by visiting the BC Data Catalogue at https://catalogue.data.gov.bc.ca/organization/data-innovation-program-dip

*jeffrey.hicks@utoronto.ca. I am grateful for thorough feedback from David Green, Antonella Mancino, and William Warburton, and have benefited from helpful conversations with Anke Kessler, Joshua Gottlieb, Samuel Gyetvay, Helen Kissel, Kory Kroft, Kevin Milligan, David MacDonald, Hugh Shiplett, Michael Smart. I also thank audience members at the Canadian Economics Association Annual Conference, the Canadian Public Economics Group Annual Conference, and the Society for Labor Economists Annual Conference for helpful comments. Ingrid Monsivais Ibarra provided excellent research assistance. I'm grateful to an anonymous reviewer at the Ministry of Social Development and Poverty Reduction. Funding came from the Social Sciences and Humanities Research Council of Canada and the UBC Public Scholars Initiative.

Administration and application processes in safety net programs are shifting away from in-person interactions to phone and Internet platforms. This shift is evident in the closure of field offices where current clients and new applicants can receive face-to-face assistance. As two examples, about 10% of US Social Security Administration offices have closed since 2000 (Deshpande and Li, 2019), and in my Canadian setting, social assistance offices saw even more widespread closures. I study how removing this in-person support affects access to transfer programs.

The programs I study are Income Assistance (IA) and Disability Assistance (DA) in British Columbia (BC), Canada, which are transfers to low-income households without and with work-limiting disabilities. IA resembles US Temporary Assistance for Needy Families (TANF), and therefore, I use the terms *welfare* and *IA* synonymously. DA resembles US Supplemental Security Income (SSI) for adults because eligibility is based on disability, irrespective of prior employment.

Field offices for these programs offer in-person information and application support. The application process can be complex, features multiple steps over weeks (or months), and requires substantial attention by applicants. Households can apply via phone and Internet platforms, but these do not entail in-person support and applicants would still need to deliver signed documents in person to government offices. Office closures, therefore, mean less in-person support and increased hassle for potential applicants. During my period of study (1999 to 2014), the number of offices was approximately halved. Anecdotally, these closures appear to have discouraged applicants (Klein, Reitsma-Street and Wallace, 2006) — and perhaps in recognition of this, some in-person support has recently been restored (see Section 1).

I estimate how these closures affected caseloads using a stacked difference-in-difference design. I use longitudinal administrative data for all BC residents (4.7 million in 2014) that captures IA/DA receipt, household structure, healthcare use, education, immigration status, and households' residence location. This rich administrative data allows measurement of caseloads across groups that vary in their health, eligibility, and familiarity with the programs. It also allows me to avoid survey-based measurement that is common in the safety-net literature (Meyer and Mittag (2019)).

In my preferred estimates, an office closure reduces combined IA and DA caseloads in the local area by 9.25% on average, equivalent to 0.45% of the working-age population (age 18-60). Reductions are higher for IA (11.50%) than for DA (1.58%), indicating that closures disproportionately affected persons with fewer medical barriers to work. Caseload declines were largest when there were no other nearby offices and, therefore, larger in more rural areas.

To tease out mechanisms, I examine whether effects were strongest among people who are plausibly less aware of the programs or are less adept at navigating the application. First, I compare the effects on young adults with and without program exposure during their teenage years via their parents. This comparison is motivated by research that finds a causal link between teenage exposure and future program receipt plausibly due to parental information transmission (De Haan and Schreiner, 2025; Hicks et al., 2023; Hartley, Lamarche and Ziliak, 2022). I find that caseload declines (in percent terms) are modestly larger for young adults *without* parental exposure, but the difference is not statistically significant. Second, I examine effects among all adults (regardless of age) separately by whether they previously received benefits themselves. Here again, effects are larger among those without prior benefit history, with differences significant at the 5% level. Both heterogeneity cuts suggest that office closures reduce take-up disproportionately among people less familiar with the programs. Third, I compare immigrants to non-immigrants, based on the conjecture that immigrants may be less aware of the programs or struggle with the application due to language barriers.[1] But I find no compelling evidence of differences in effect size.

With imperfect screening technology, these caseload declines could be among “truly" eligible or ineligible applicants (Kleven and Kopczuk, 2011). To *indirectly* explore this, I examine whether caseload declines were larger among plausibly healthier persons, based on the intuition that welfare and disability assistance are intended for people with temporary or permanent barriers to work. This is an indirect exercise because health is an imperfect proxy for eligibility — particularly for welfare, which does not condition eligibility on health status, even though it nonetheless

[1] The premise that immigrants are less able to navigate the application process or are less aware of the programs is just a conjecture. Refugees (a small subset of immigrants) have high usage rates of social assistance (Frenette, Lu and Schellenberg, 2015), but this at least partly reflects their proneness to deep poverty.

insures against health shocks (Hicks, 2023). I find that the *welfare* caseload declines are probably larger among healthier individuals, as measured by total healthcare utilization and diagnosed mental illness. However, this pattern is heavily driven by age: caseload declines are larger among younger persons who tend to be healthier. Aside from the health/age heterogeneity, I find similar declines across family types, specifically childless adults versus parents. There are even welfare caseload declines among mothers of pre-school-age kids, which is notable because they have fewer eligibility requirements and arguably more barriers to work.

Finally, building on approaches used by Castell et al. (2024) and Finkelstein and Notowidigdo (2019) for measuring targeting, I find that households who no longer access assistance due to closures have the same counterfactual duration of benefit receipt as do households who take up benefits despite a closure (under potentially strong assumptions discussed in Section 4.2).

Literature and Contribution: Deshpande and Li (2019) find that Social Security Administration office closures reduced enrollment in Social Security Disability Insurance (SSDI) and SSI by 15% and 18% respectively, while Rossin-Slater (2013) finds that office availability improves take-up of the Special Supplemental Nutrition Program for Women, Infants, and Children by 6%. My results qualitatively support those findings, but I find smaller effects on disability assistance take-up than found for SSI (the most comparable US analog).[2] To my knowledge, I provide the first estimates of office closures' effects on welfare enrollment, the Canadian version of which is most analogous to TANF, and to a lesser extent SNAP. My rich data linkage also allows me to investigate multiple dimensions of targeting and plausible mechanisms.

Moreover, in contrast to much of the literature, I can compare the effect of the same treatment (office closures) on different programs (welfare vs. disability assistance) in the same setting, since the same administrative structure serves both programs. Concurrent work by Wu and Meyer (2023) has the same advantage: they compare take-up reductions across a range of safety net programs in Indiana when caseworker assistance shifted from in-person to Internet- and phone-based platforms.

[2] Relatedly, Foote and Rennane (2019) find that a major improvement of iClaim, the online application portal for DI and retirement benefits in the US, increased applications and benefit receipt. Zuo and Powell (2023) find that broadband access increases SSDI and SSI take-up rates.

Like my results, their largest effects are found for welfare enrollment (TANF).

My results are also consistent with experimental studies that raised take-up of government benefits by providing information and application assistance (Finkelstein and Notowidigdo, 2019; Castell et al., 2024; Hermes et al., 2024). For example, Castell et al. (2024) find that assigning a social support worker to assist in the application increased take-up by 31%. Researchers have also documented take-up effects of specific aspects of program administration – notably, intake and recertification interviews (Giannella et al., 2023; Homonoff and Somerville, 2021).

The normative implications depend on who is screened out (Nichols and Zeckhauser, 1982; Besley and Coate, 1992; Diamond and Sheshinski, 1995; Kleven and Kopczuk, 2011; Rafkin, Solomon and Soltas, 2023) and, therefore, the empirical literature has emphasized the targeting properties of related interventions. For example, both Castell et al. (2024) and Finkelstein and Notowidigdo (2019) find that application assistance disproportionately increases take-up among less eligible households, while in contrast, Deshpande and Li (2019) find that office closures worsen targeting in SSDI and SSI. Overall, I find limited change in the proportional targeting of welfare benefits, although caseload declines were modestly larger among younger and healthier adults.

1 Institutional Details

British Columbia's Income Assistance (IA) and Disability Assistance (DA) system is the primary income support program available to very low income households. IA provides support to those without significant work-limiting disabilities, while DA supports persons with such disabilities.[3]

Eligibility: Households with very low income and assets are eligible for IA. Most welfare recipients are required to search for and accept work[4] or undergo training. Earnings received while on welfare reduce benefits dollar-for-dollar (in some years, there were small earnings exemptions). Starting in 2002, most new welfare applicants had to demonstrate some degree of financial indepen-

[3]A severe mental or physical impairment that directly and significantly restricts the person's ability to perform daily living activities, that requires either an assistive device or significant help from another person or service animal, and that is expected to persist for more than two years.

[4]Mothers were exempt if their youngest child was less than age 7 between 1996 and 2001, or age 3 after 2001.

dence over the prior two years.[5] Both childless adults and households with children are eligible. To be eligible for DA, a person must also obtain a certification of long-term and severe work-limiting disability, but is exempt from work search requirements and the financial independence test.

Benefit Amounts: Benefit amounts are determined by the number of household members and the shelter expenses incurred by the household. Benefits are unrelated to prior employment history. Figure A.1 shows that nominal monthly benefit rates are almost flat over time (not indexed to inflation) during my study period. For IA, single childless households received \$500 to \$600 per month while single parents received approximately \$900 per month. IA benefits are lower than DA benefits — single childless adults received \$800 to \$900 if receiving DA. For all groups, these benefit rates are below the poverty line.[6] IA and DA recipients may also receive in-kind transfers, such as for uninsured medical devices, and full coverage for medications which are not universally insured for non-IA/DA recipients (see Hicks (2023) for more details). The other major income support for households in Canada, particularly those with children, are benefits administered through the tax system irrespective of IA/DA eligibility.[7]

Long-Run Trends: Welfare reform in 2002 tightened the eligibility criteria for welfare, increased work search requirements and enforcement, and intensified application burdens, all of which collectively reduced IA caseloads. See Green et al. (2021) for details and Figure A.2 for long-run caseload trends. DA is the larger program today, partly due to declining IA caseloads and partly due to rising DA caseloads (Kneebone and White, 2014).

Application Procedures

The application for IA contains multiple stages that are visualized in Figures A.3 and A.4.

[5] Typically defined as having market income $\geq$ \$7,000 in each of the prior two years.

[6] Statistics Canada's before-tax low-income cut-offs for cities with over 500,000 inhabitants in 2001 was \$1,972 per month for a family of two (see Table 11- 10-0241-01).

[7] These have changed over time and include the National Child Benefit, the Universal Child Care Benefit (UCCB), and the Child Tax Benefit. Generally, these are tax benefits given to households with children with far less means-testing (or no means-testing for the UCCB) and lower dollar amounts. Hicks et al. (2023) show that, in their sample of IA recipient families, these tax benefits were about 25% of annual disposable income, while IA income was over 50%.

Pre-Application Stage: In the *pre-application* stage, applicants undergo a *non-binding* initial assessment of their eligibility either in-person, by phone, or via an online self-serve portal (phased-in since 2010). When conducted by phone or in-person, ministry staff could, throughout the process, ask whether the applicant wishes to continue, effectively granting staff the discretion to discourage would-be applicants (Ronayne et al., 2009). The process after the pre-application varied over time.

Application Stage: Before 2002, the applicant then had to collect supporting documents used for eligibility determination and undergo an intake interview, typically in person. Based on the documentation and interview, an Intake Worker assessed the applicant's eligibility. Starting in 2002, a three-week waiting period was introduced, during which applicants had to undergo a 14-day self-directed work search.[8] If evidence of work search was deemed unsatisfactory, or the applicant did not report back 14 days after the initial contact, staff could terminate the application. After completion of the work search and documentation steps, applicants advanced to the interview.

Documentation: The documentation requirements could be numerous. As of 2008, up to thirty forms could be required, some of which had to be submitted in person (Ronayne et al., 2009).[9] The delivery of these documents (to validate identities) could be done at other government offices, not just offices serving IA/DA.[10]

Additional Procedures for Disability Assistance: To apply for disability status, the individual must find a physician to fill out a thirty-page booklet describing the applicant's work-limiting disability. Unlike for welfare applications, there was no three-week waiting period but application processing times were longer. Currently, the government targets a maximum of 45 days.

[8]Exemptions to this waiting period were available for those experiencing "undue hardship", such as lacking food or shelter or fleeing domestic abuse. Single mothers with children under age three also eventually gained an exemption.

[9]An Ombusperson report (Ronayne et al., 2009) documented the experience of an applicant suffering from major depressive disorder and Hepatitis B and C. This person was required to submit the following: a recent electricity bill, driver's license, notice of unpaid strata fees, confirmation of monthly strata fees, a property tax assessment, a letter confirming unpaid property taxes, a Registered Retirement Savings Plan (RRSP) statement, a letter from the bank confirming that the $2,600 in the RRSP could not be withdrawn, a letter confirming tax owed to the Canada Revenue Agency, two tax return Notice of Assessments, a bank statement, credit card, and mortgage statements, a doctor's note, a passport, a Record of Employment, estate documents from a sibling, and a Social Insurance card. Only after 2008 were staff required to provide a document checklist to applicants with associated deadlines.

[10]Such as Service BC offices which offer other services such as driver's licenses (Ronayne et al., 2009).

The Role of Offices: IA and DA are served by the same field offices. Office staff provided information to current and prospective clients, conducted pre-application and intake interviews, decided which supporting documents were required, and ruled on applicants' eligibility. They were responsible for assisting clients in filling out forms, although there were no performance incentives to encourage this assistance. More recently, when the application incorporated a mandatory online orientation, offices provided some computer access. Notably, offices did not conduct active outreach to inform potential beneficiaries about programs and eligibility. The administrative functions of offices slowly shifted to online and call-center platforms starting in 2006 when the telephone system gained prominence, and then, starting in 2010, the Internet portal. Simultaneously, the number of offices declined from 128 to 64 (sources described below). Aside from anecdotal reports (Klein, Reitsma-Street and Wallace, 2006; Ronayne et al., 2009; Biscoe et al., 2018), there is no evidence of how the closures affected caseloads.[11]

More recently, starting after 2014, the province began partnering with other government offices that administer services such as driver's licenses, particularly in rural communities, to train their staff to provide support to IA/DA applicants. Additionally, starting in 2019, an outreach program was established to support vulnerable individuals (*e.g.*, unhoused persons) in connecting with benefits. These expansions of service generally fall outside the time of my study.

2 Data

My empirical strategy relies on three data inputs: (1) caseloads in a given geographic area and the total population living there; (2) a list of offices in operation and their locations for each year; and (3) characteristics of the households. I describe the sources for these below.

Caseload Data: I access caseload data for the universe of IA and DA recipients from 1991 to 2018, containing monthly records of benefit receipt for each individual, the family members on

[11] Interviews conducted as part of a governmental commission (Green, Kesselman and Tedds, 2021) indicated that public library workers and political constituency offices (personal communication, September 25th, 2019) sometimes provided *de facto* support for those trying to use computers to complete the application.

each claim, and the location of residence. I only observe caseloads, not *applications*.[12] Recorded benefit amounts reflect regular IA payments, not in-kind transfers such as the drug subsidy.

Population Data: During the study period, every resident of the province was legally required to register with the provincial universal health insurance program (Medical Services Plan (MSP)). Therefore, the MSP registry is my primary source for population counts by area. A very small share of residents may not appear in the MSP registry, but population counts from the registry are extremely close to population estimates from Statistics Canada.

Demographics:

Age and Household Type: The MSP registry allows me to classify individuals according to their household status (*i.e.*, childless adult vs parents) and age. In rare cases when the household is not in the MSP registry, I use the household type recorded in the IA/DA records.

Immigrant Definition One: I classify someone as recently immigrated if they appeared in the MSP registry for the first time within the past five years, which captures both international and inter-provincial migrants. The MSP registry begins in 1985, so persons coded as immigrants could have lived in the province before 1985. My first office closure occurred in 1999, so immigrants in my sample will, at a minimum, have been outside the province for a decade.

Immigrant Definition Two: An alternative approach strictly captures international immigrants for a narrower population subsample: for mothers who gave birth in BC, I observe the mothers' own location of birth. For example, a mother born in the US but who gave birth in BC will be flagged as an international immigrant. The birth records begin in 1985.

Prior Program Exposure: I will split young adults (age 20-24) based on whether their household received IA/DA during their teenage years (age 14-17) using the longitudinal aspect of the data.[13] More generally, in any given year, I will split adults into whether they had ever received IA/DA

[12] Hicks (2023) accessed application files but these were not made available for other research projects.

[13] The age 20 cutoff follows Hicks et al. (2023)'s baseline cutoff for "young adults" when examining intergenerational welfare transmission, but this threshold could be lowered without qualitatively changing results. The age 14 cutoff for teenage exposure is dictated by the data starting in 1990.

before that year as an adult (age 18+). Both splits are intended to proxy for awareness of the programs – one through parental exposure, one through personal experience.

Educational Attainment: Among those individuals that appeared in the primary and secondary school enrollment records in 1991 or later, I will classify individuals according to whether they graduated with a high school diploma (the enrollment records begin in 1991).[14]

Health Records: Medically-necessary care is fully publicly-insured. I access all outpatient billings and all inpatient and day-surgery hospital visits, and the associated diagnosis codes for a given treatment/visit.[15] The main exclusions from these data are non-universally insured care: pharmaceuticals, dental, and allied health professionals (e.g., chiropractors).

Defining Geography: The administrative files contain the Forward Sortation Area (FSA) of the household's residence, which is my definition of an “area". An FSA is the first three digits of a postal code, similar to the five-digit ZIP code in the US. Urban FSAs are geographically narrow, rural ones are larger. There were 191 FSAs in 2011 with a median total population of 20,399[16]. Figure B.1 maps their boundaries, which were quite stable over the sample period.

Office Locations and Closures: I collect address information for each office from 1991 to 2017 using archived government sources described in Appendix B. I identify office closures in two ways. I first compile a list of publicly announced closures from archival sources (also detailed in Appendix B). Publicly announced closures are, however, an incomplete list since most were not publicized. I infer non-publicized closures by comparing year-over-year changes in the directory of operating offices. If an office disappears from year $t-1$ to year t, I infer it as closing in year t. I exclude simple address changes or office amalgamations. See Appendix B for details. Because the directories were accurate as of July of that year, the closure could have been in year $t-1$ or t.

Assigning Areas to Offices: For each of the 100,000 six-digit postal codes, in each year, I identify the nearest office based on straight-line distance between the postal code's centroid and the office's

[14] In BC terminology, a “BC Secondary School Graduation" or "Dogwood" diploma.

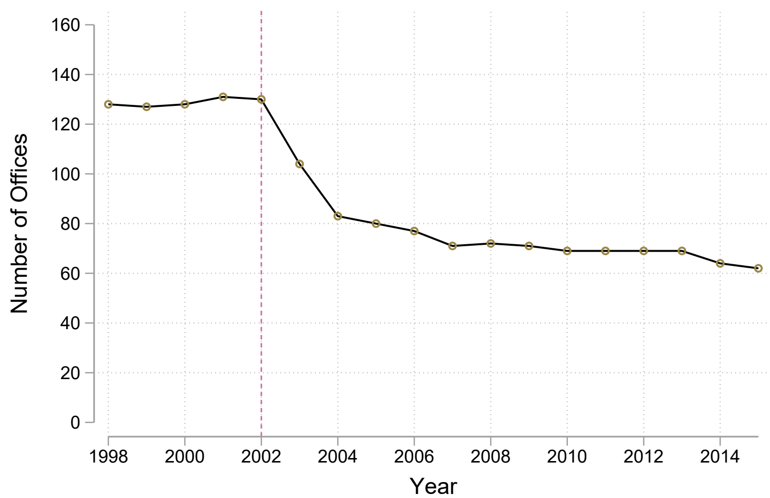
[15] The outpatient billings and hospital inpatient records indirectly capture about 75% of Emergency Room (ER) visits (Peterson et al., 2021). The full ER data is not available for the study period.

[16] Source: http://datacentre.chass.utoronto.ca/census/index.html

location. I consider a six-digit postal code as exposed to a closure if its nearest office closes. Then, for each three-digit postal code (FSA), I determine the fraction of the FSA's population exposed to a closure based on the set of exposed six-digit codes. Appendix B contains details.

Trends in Office Locations: Figure 1 plots the number of offices in each year between 1998 to 2015. During this time, the number of offices decreased from 128 to 64. In small rural communities, the initial wave of closures left them without an office (illustrated in Figure B.2). However, closures occurred in urban areas as well, as shown in Table B.1. For example, the number of offices in Vancouver declined from 21 to 7.

Figure 1: **Number of Offices in Operation by Year**



The number of IA/DA offices in operation in each year. Series construction and sources are described in Section 6. The same office serves both IA and DA.

The government did not explain how closure locations were chosen. Table B.1 and Figure B.2 show that closures were fairly dispersed across community size and the province. In the following section, I examine the area characteristics that predicted closures.

3 Empirical Strategy

I use a stacked difference-in-difference strategy that, within a given time frame, compares trends in areas that experienced closures against areas that did not. This approach avoids problems with two-way fixed effects estimators in staggered designs (see Goodman-Bacon (2021) and Roth et al. (2023) for overviews). Below, I outline each aspect of identification and estimation.

Stacking: The stacked difference-in-difference compares areas (g) that experienced a closure in year t ($D_{g,t} = 1$) against areas that did not experience a closure any time between $t-4$ and $t+4$ ($D_{g,t} = 0$). To illustrate the *stacking* of events, consider a difference-in-difference for a single event year t, given by $y_{s,g} = \alpha_s + \alpha_g + \beta_t \mathbb{1}\{s > 0\} D_{g,t} + \epsilon_{g,s}$ where s is the year relative to t. β_t represents the effect of year t closures. A β_t can be estimated for each set of closures, then averaged across all closure years t. The stacked difference-in-difference does this by creating a dataset for each t whose unit of observation is (g, s), appending the datasets together, and estimating a single difference-in-difference in event time s on the combined dataset. See Deshpande and Li (2019) and Cengiz et al. (2019) for examples.

Specification: I model the conditional expectation of the outcomes in exponential form:

$$E\left(\frac{y_{s,g,t}}{p_{s,g,t}}\right) = \exp\left(\alpha_{g,t} + \alpha_{s,t} + \beta \mathbb{1}\{s > 0\} D_{g,t}\right) \tag{1}$$

Where the outcome $y_{s,g,t}$, such as the number of benefit recipients, is normalized by the area's population $p_{s,g,t}$. $\alpha_{g,t}$ and $\alpha_{s,t}$ are fixed effects for area and event time, respectively, interacted with event year t. Note that $\alpha_{g,t}$ are perfectly collinear with $D_{g,t}$. β is the percent change in the conditional expectation (Roth and Chen, 2023):

$$\beta = ln\left(\frac{E(\frac{y_{s,g,t}}{p_{s,g,t}} \mid D_{g,t} = 1, s \geq 0)}{E(\frac{y_{s,g,t}}{p_{s,g,t}} \mid D_{g,t} = 1, s < 0)}\right) - ln\left(\frac{E(\frac{y_{s,g,t}}{p_{s,g,t}} \mid D_{g,t} = 0, s \geq 0)}{E(\frac{y_{s,g,t}}{p_{s,g,t}} \mid D_{g,t} = 0, s < 0)}\right) \tag{2}$$

Estimating percent changes is useful for comparing effects across sub-groups of the population that differ in their baseline benefit usage. If β is larger in one group than in another, I say that closures affect the proportional targeting of the program between those groups. Modeling the conditional mean as exponential rather than log-linear also allows for proper treatment of zeros in the outcome variable and other potential issues in log-linear approximations (see Roth and Chen

(2023) and Silva and Tenreyro (2006) for overviews).

Outcome Measurement: I calculate the outcome $\frac{y_{s,g,t}}{p_{s,g,t}}$ using the working-age population (age 18-60). Children (< 18) and seniors (> 60) are not used to construct the numerator or denominator.[17]

Assumptions: Given the exponential model, for β to be causally interpreted as an office closure effect, the growth rate of $\frac{y_{s,g,t}}{p_{s,g,t}}$ in treated areas must trend parallel to the growth rate in control areas (Wooldridge, 2023). To assess parallel *pre*-trends, I estimate:

$$E\left(\frac{y_{s,g,t}}{p_{s,g,t}}\right) = \exp\left(\alpha_{g,t} + \alpha_{s,t} + \sum_{h\neq -2} \beta_s \mathbb{1}\{s = h\} D_{g,t}\right) \tag{3}$$

Where the base period is $s = -2$ to account for the closure occurring between $s = -1$ and $s = 0$.

The most plausible potential violation is that many closures occurred around the time of welfare reform. If reform differentially affected treated areas, β would capture this effect. Two exercises will assuage this concern: First, as described below, I use a subset of control areas that are observably similar to treated areas (“matching"). Second, I will show similarly-sized treatment effects for closures that occurred in years outside of welfare reform.

A subtler potential complication is spillovers from treated to control areas. When an office closes, some applicants may be diverted to offices that serve control areas, thereby amplifying office congestion that deters potential applicants residing in control areas. This spillover would imply that the estimate of β is a lower bound on the true effect of closures in treated areas. In practice, only a small fraction of control areas in the analysis sample neighbor the treated areas, and when removing these remaining nearby control areas, the results are almost identical.

Defining Treatment and the Sample of Areas: I define an area g as experiencing a closure in year t ($D_{g,t} = 1$) if more than 20% of its population was exposed to a closure. Control areas are those with 0% exposure. Areas with exposure $\in (0, 20\%]$ are dropped.[18] I exclude all areas that

[17] Kids typically become eligible to apply for benefits as their own household (separate from parents) at age 19, although there are exceptions that allow earlier application.

[18] Dropping the areas with positive but very small exposure ensures that the “closure event" is ex-ante sufficiently meaningful to have a statistically detectable effect on area-wide caseloads.

experienced a closure in years other than t within the event window ($s \in [-4, 4]$) to eliminate confounding multiple events. I exclude the very few areas that had major boundary changes within the event window.[19]

Estimation and Inference: Equation 1 is estimated using Poisson quasi-maximum likelihood (implemented by Correia, Guimarães and Zylkin (2020)) which tends to be the most robust and asymptotically efficient estimator for exponential models (Gourieroux, Monfort and Trognon, 1984).[20] Standard errors are clustered at the area level (g) to account for correlation over time within areas.

3.1 Regression Sample, Matching, and Weighting

The first three columns of Table 1 show the number of treated and control areas in each event year t, denoted $N_{d,t}$ and $N_{c,t}$ respectively. In total, there are 122 instances of an eligible area experiencing a closure and 1,214 eligible control areas. Treated areas outnumber closures because some offices served multiple areas.

What predicts office closures? The first panel of Table 2 shows the average characteristics of treated and control areas in pre-closure periods. Treated areas had lower rates of IA/DA receipt and were more likely to have another office nearby (as measured by the average distance to the second-nearest office). Treated areas also had slightly smaller shares of young adults and single parents, two groups that rely more heavily on assistance. These facts suggest the government tended to close offices in areas with lower demand for assistance and more nearby substitute offices.

The slightly different demographics between treated and control areas could cause point-in-time shocks – such as 2002 welfare reform or the 2009 recession – to differentially affect treated and control areas. To alleviate this concern, I use a matched control group.

Matching: For each event year t dataset, I use Coarsened Exact Matching (Blackwell et al., 2009) to match treated and control areas based on the pre-closure averages of: the prevalence

[19] I consider a boundary change to occur when the area's population changes by 25% or more year-over-year.

[20] Asymptotically most efficient if the mean-to-variance ratio is constant and more robust to distributional misspecification than other estimators (Gourieroux, Monfort and Trognon, 1984; Silva and Winkelmann, 2024)

Table 1: **Treated and Control Areas for Each Event Year**

Year	Full Sample			Matched Sample		
	Control Areas ($N_{c,t}$)	Treated Areas ($N_{d,t}$)	# Closures	Control Area ($N_{c,t}$)	Treated Areas ($N_{d,t}$)	# Closures
1999	93	4	1	9	2	1
2000	91	3	2	7	3	2
2001	83	5	5	10	4	4
2002	75	15	13	34	14	12
2003	62	36	15	48	31	15
2004	83	2	1	14	2	1
2005	85	1	1	3	1	1
2006	85	20	4	50	18	3
2008	142	2	1	18	2	1
2009	129	6	2	29	5	2
2013	137	22	5	75	22	5
2014	149	6	2	14	6	2
Total	$N_c = 1214$	$N_d = 122$	52	$N_c = 311$	$N_d = 110$	49

Note: The first three columns show the number of treated and control areas for each event year from the full sample described in the text, and the corresponding number of office closures. There are fewer office closures than treated areas because some offices served multiple areas. The second three columns show the same counts after matching control areas to treated areas based on pre-event characteristics, as described in the main text.

of young adults and single parents, total population, the fraction of residents receiving IA/DA, and the growth of the IA/DA caseload. Each matching variable is split into two bins at the median. Observations are put into groups (indexed by h) based on the interaction of the bins; *e.g.*, with two covariates each split into two bins, there are four groups. Observations are then re-weighted based on their group. Treated areas have a weight equal to one. Control areas have weight $\frac{N^h_{d,t}}{N_{d,t}} / \frac{N^h_{c,t}}{N_{c,t}}$, where $N^h_{c,t}$ and $N^h_{d,t}$ are the number of control and treated areas in group h. Groups with $N^h_{d,t} = 0$ or $N^h_{c,t} = 0$ are dropped.

As shown in Table 1, the matched control group retains 311 out of 1,214 eligible control areas. The second panel of Table 2 shows the average area characteristics after matching — by definition, characteristics become significantly more balanced.

Estimation Weights: Observations are weighted by the area's average pre-closure population ($p_{g,t,s<-1}$), similar to Deshpande and Li (2019). Additionally, following Wing, Freedman and Hollingsworth (2024)'s advice for stacked difference-in-difference estimators, observations from event year t receive weight $\frac{N_{d,t}}{N_d} / \frac{N_{c,t}}{N_c}$. This ensures that the treatment effect in an event year dataset

Table 2: **Average Pre-Closure Characteristics of Treated and Control Areas**

	Treated Mean	**Control Mean**	**Difference**	ρ
		Full Sample		
Matching Variables				
Fraction Single Parent	0.049	0.057	-0.008	0.000
Fraction 18 to 25	0.164	0.173	-0.008	0.002
Population	15409	13175	2234	0.010
% Change in Caseload Per Capita	-0.005	-0.006	0.000	0.559
Caseload Per Capita	0.062	0.079	-0.017	0.001
Unmatched Variables				
Population Growth	0.015	0.010	0.005	0.030
Distance to Nearest Office (km)	5.45	10.10	-4.65	0.006
Distance to Second Nearest Office (km)	8.32	25.93	-17.61	0.000
		Matched Sample		
Matching Variables				
Fraction Single Parent	0.050	0.054	-0.004	0.094
Fraction 18 to 25	0.166	0.168	-0.002	0.531
Population	14757	14343	414	0.666
% Change in Caseload Per Capita	-0.004	-0.005	0.001	0.063
Caseload Per Capita	0.062	0.068	-0.006	0.391
Unmatched Variables				
Population Growth	0.014	0.011	0.003	0.267
Distance to Nearest Office (km)	4.51	7.94	-3.44	0.139
Distance to Second Nearest Office (km)	8.18	22.49	-14.31	0.011

Note: This table shows mean characteristics of treated and control areas using pre-closure observations. For a given area characteristic in event year y, $X_{g,y}$, I regress $X_{g,y}$ on fixed effects for y and on the treated dummy. The treated mean column is the constant plus the coefficient on the treated dummy. The control mean column is just the constant. The third column shows the coefficient on treated dummy, and ρ is the p-value for the test that it is 0. The first panel uses the full sample of areas, the second panel shows uses the matched sample (described in Section 3.1).

receives weight proportional to that event year's share of overall treated units. After combining these weights with the within-event-year matching weights, the final weight reduces to $p_{g,t,s<-1} \frac{N^h_{d,t}}{N_d} / \frac{N^h_{c,t}}{N_c}$ for control units and $p_{g,t,s<-1} \frac{N_{d,t}}{N_d} / \frac{N_{c,t}}{N_c}$ in treated units.[21] I will show robustness to alternative weights.

4 Results

4.1 Main Results

Panel (a) of Figure 2 shows the main event study estimates from equation 3 using the matched control group, and, to illustrate the role of matching, estimates using the full control group. Pre-

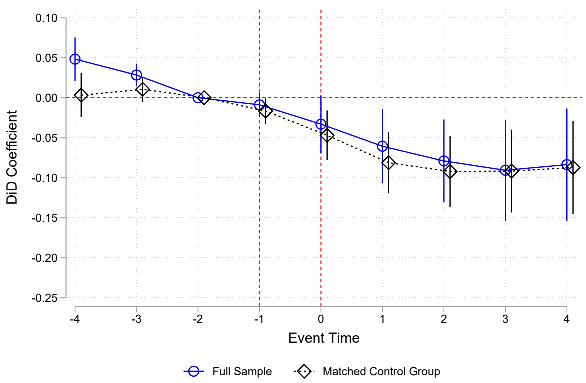
[21] Wing, Freedman and Hollingsworth (2024) derive these weights in a linear model, rather than the exponential model I use, but the intuition generalizes. I apply the weights to the observations likelihood function contribution.

trends are not parallel when using the full control group, but are parallel using the matched control. I show results using the matched control hereafter.

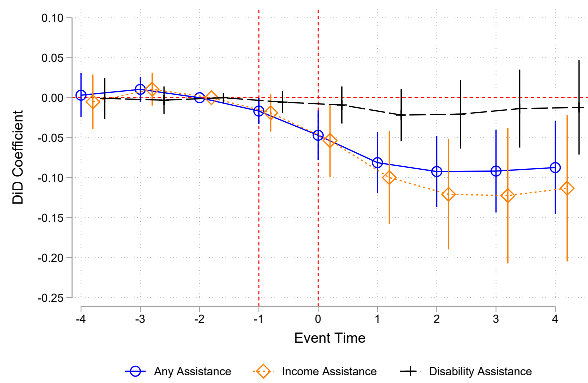
A very small effect materializes in $s = -1$ followed by a larger effect in $s = 0$ and $s = 1$. These dynamics are consistent with (a) the closure occurring sometime between $s = -1$ and $s = 0$, and (b) closures reducing the net-inflow of recipients, the effect of which accumulates over the initial post-closure years.

Figure 2: **Baseline Event Study**

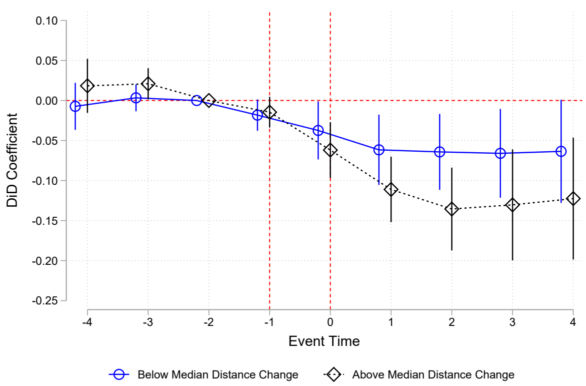
(a) **Using Full Control vs. Matched Control**



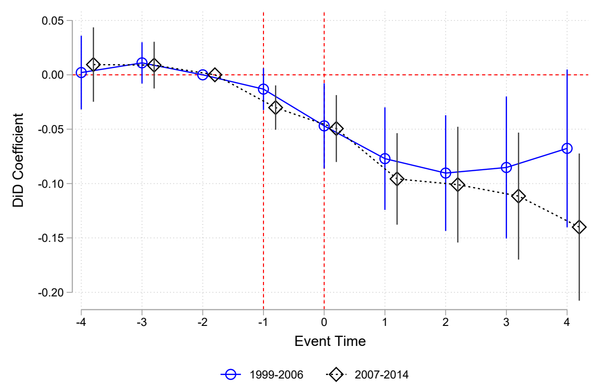
(b) **Comparison of Program Type**



(c) **By Change in Distance to Nearest Office**



(d) **By Time Period of Closure**



Note: Estimated β_s from equation 3 for each s from -4 to 4, and 95% confidence intervals. The red vertical lines denote the office closure time, sometime between -1 and 0. Panel (a) shows estimates using the full control and matched control. Remaining panels use the matched control. Panel (b) shows effects on Income Assistance and Disability Assistance separately, in addition to combined. Panel (c) shows estimates based on the median distance change. Panel (d) shows estimates from the set of closures that occurred between 1999 and 2006 and 2007 to 2014, respectively.

Table 3 shows estimates of β from equation 1 alongside the implied effect on $E\left(\frac{y_{s,g,t}}{p_{s,g,t}}\right)$, which I call the Level Effect. Results are shown for the combined IA and DA caseload and separately for each program. The first three columns measure caseloads ($y_{i,t}$) by the number of adult recipients during the year. Estimates indicate that a closure reduces caseloads, on average, by 9.25%, which is equivalent to 0.45% of the working-age population. The effect on IA caseloads is 11.50% while the effect on DA is a statistically insignificant 1.58%. The last three columns measure caseloads ($y_{s,g,t}$) by total months of benefits received rather than persons. Results are qualitatively unchanged.

Assuming the effect on aggregate benefit months is driven only by changes in the number of recipients, dividing the former by the latter equals the average lost months per would-be recipient. For example, closures reduce the fraction of residents that are recipients by 0.0045 and reduce benefit months per resident by 0.04. Therefore, persons who no longer received assistance because of closures would have recieved 8.88 ($\frac{0.04}{0.0045}$) benefit months per year. Similarly, they would have received \$7,171 ($\frac{32.27}{0.0045}$) per year, or \$807 per month ($\frac{7{,}171}{8.88}$), which is about the mean benefit rate shown in Figure A.1.

Table 3: **Stacked Difference-in-Difference Estimates of Closures' Average Effect**

	Recipients Per Capita			Benefit Months Per Capita		
	Any IA/DA	IA	DA	Any IA/DA	IA	DA
Treated × Post	-0.0925***	-0.115***	-0.0158	-0.0892***	-0.132***	-0.0166
	(0.0279)	(0.0422)	(0.0270)	(0.0269)	(0.0464)	(0.0273)
Outcome Mean	0.0688	0.0513	0.0197	0.599	0.390	0.209
Level Effect	-0.0045	-0.0030	-0.00037	-0.040	-0.026	-0.0042
Level Effect SE	0.0014	0.0011	0.00064	0.012	0.0092	0.0069
Observations	2947	2947	2947	2947	2947	2947

Note: This table shows estimates of β from equation 1 and standard errors. The *Outcome Mean* is the average $\frac{y_{s,g,t}}{p_{s,g,t}}$ among treated areas in the pre-closure period. The first three columns define $y_{s,g,t}$ as the number of recipients. The last three define $y_{s,g,t}$ as the aggregate number of benefit months paid out. The *Level Effect* reports the effect on $E(\frac{y_{s,g,t}}{p_{s,g,t}})$ rather than on $ln(E(\frac{y_{s,g,t}}{p_{s,g,t}}))$, using *pre-closure* averages in the treated group, with the standard error calculated using the delta method. Event time -1 and 0 are excluded as these are partial treatment years. *** $\rho < .01$, ** $\rho < .05$, * $\rho < .1$

Potential Congestion Spillover: As discussed above, if residents of treated areas divert to offices in control areas, this could create congestion that could lower take-up among control area residents, which would bias β toward zero. In column (2) of Table 4, I drop control areas that are near treated areas (about 10% of the sample).[22] The estimate of β only marignally rises (in absolute terms).

Continuous Measure of Exposure: Rather than using a binary treatment variable, I could define a continuous treatment as the share of an area's population exposed to a closure (as defined in Section 2), where this share is zero for the control areas. Intuitively, the drop in an area's caseloads should be larger when more of the area's population is affected. Column (3) of Table 4 shows the results using this approach. The estimate of β is larger but qualitatively similar. Figure C.1 extends this logic by plotting the relationship between the share of the population exposed and the area's change in caseloads following the closure, just for treated areas. The relationship is negative and linear, consistent with predictions. I use the binary specification as my baseline because the continuous model requires more assumptions for the estimator to be easily interpreted (Callaway, Goodman-Bacon and Sant'Anna, 2024) and because I measure population exposure with error.

Heterogeneity by Change in Distance to Nearest Office: If travel time to an office is a major barrier, then closures that occur in areas without nearby substitute offices should have larger impacts on caseloads. To test this, for each area experiencing a closure, I calculate the average *change* in distance between households and their nearest office following the closure. When there are no nearby substitutes, this change will be large. I split the treatment areas into those above and below the median distance change within each event year, then I estimate equation 3 separately for each subgroup while using the same control group for both. The results are shown in panel (c) of Figure 2 and columns (4) and (5) of Table 4. The effect of closures is doubled when there are no nearby substitutes (13.80% versus 6.24%), suggesting that distance to an office is a key factor.

Effects Across Event Years: The results thus far have been average effects across all event years. In panel (d) of Figure 2 and columns (6) and (7) of Table 4, I instead show estimates separately for closures that occurred between 1999-2006 and 2007-2014. There are similarly-sized effects in both sub-periods. This assuages concerns that welfare reform in 2002 is confounding the estimate.

Next, I go further by showing estimates of β from each event year dataset separately, denoted

[22] I define "nearby" controls as those where at least 20% of the population faced the closure of the *second-nearest* office in year t.

Table 4: **Extensions**

	Baseline	Drop Nearby Controls	Continuous Treatment	By Distance Change		By Closure Year	
				< Median	> Median	1999-2006	2007-2014
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated × Post	-0.0925***	-0.0940***		-0.0624**	-0.138***	-0.0844**	-0.118***
	(0.0279)	(0.0296)		(0.0300)	(0.0318)	(0.0341)	(0.0311)
Share × Post			-0.138***				
			(0.0404)				
N	2947	2660	2947	2576	2548	1750	1197

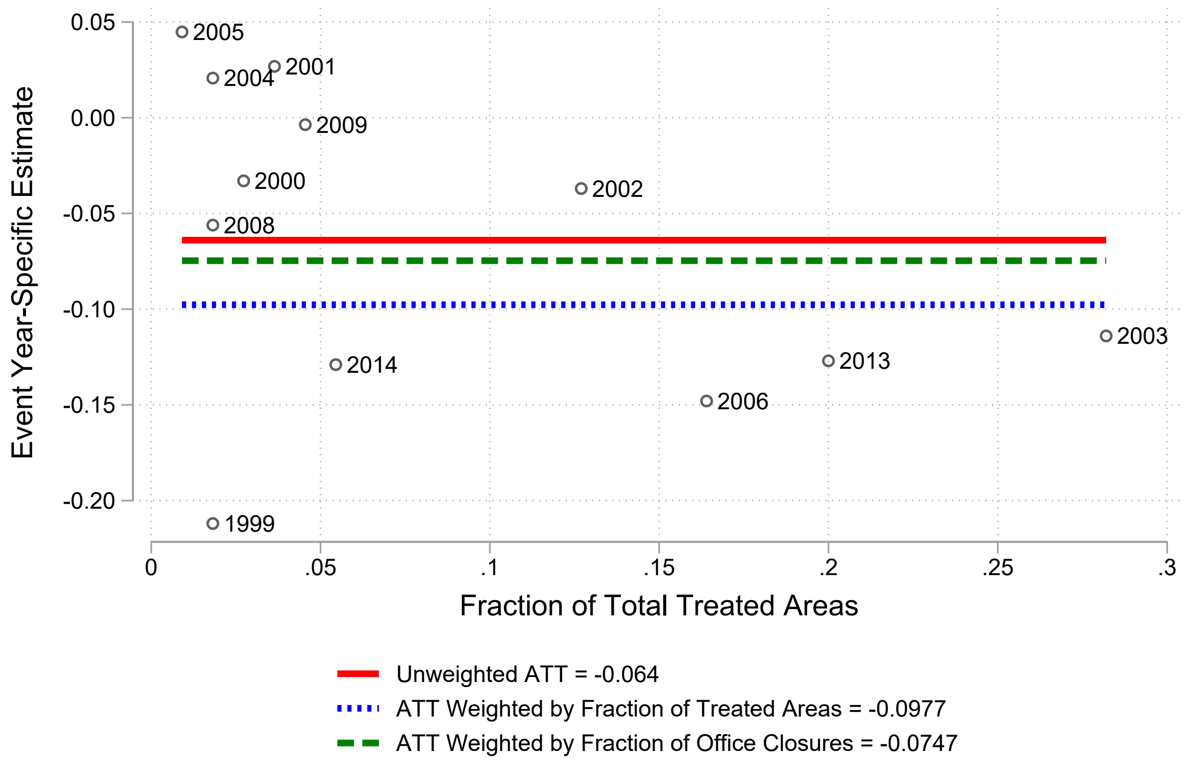
Note: This table shows estimates of β from equation 1 and standard error. The first column shows the baseline, subsequent columns examine specification changes and heterogeneity. Column (2) drops control units that were nearby treated units to test for congestion bias. Column (3) uses a continuous treatment, rather than binary treatment, defined as the fraction of the population inferred as exposed to a closure ("Share"). Columns (4) and (5) split the treatment group based on whether the closure event had a large or small effect on the distance to the nearest office — the control group is the same in both columns. Columns (6) and (7) split the sample based on the year of the closure event. Event time -1 and 0 are excluded as these are partial treatment years. *** $\rho < .01$, ** $\rho < .05$, * $\rho < .1$

by β_t. This illustrates how the average effect β is implicitly a weighted average of β_t across all event years (t), and how β changes when the weights change. Figure 3 plots the results: $\beta_t < 0$ in 9 out of 12 event years. The three horizontal lines show the weighted average of the β_t's across all t using different weights. The unweighted average is -0.064 — but Wing, Freedman and Hollingsworth (2024) argue that the natural parameter to consider is $\sum_t \frac{N_{d,t}}{N_d}\beta_t$ such that each β_t is weighted by event year t's share of total treated areas in the combined dataset. These weights are shown on the horizontal axis. The weighted average β_t using them is -0.0977, very similar to -0.0925 from the baseline stacked regression in Table 3.[23] The third horizontal line uses weights that, instead, correspond to the share of actual office closures (rather than treated areas). Using these weights, the average across β_t is -0.0747.

Taking Stock: Income Assistance (i.e. welfare) caseloads decline by 11.15% following an office closure, on average, with no detectable effect on Disability Assistance caseloads. Estimates are (i) robust to different specification choices, (ii) larger in areas that did not have nearby substitute offices, and (iii) present in most closure years.

[23]The similarity is by design, given the weighting in the baseline estimation described in the previous section.

Figure 3: **Treatment Effects on Combined IA/DA Caseload in Each Closure Event Year**



Note: The vertical axis shows estimates of β from equation 1 separately for each event year — that is, each dot shows the estimated treatment effect for closures that occurred in a given year. The dots' labels indicate which year. The horizontal axis is $\frac{N_{d,t}}{N_d}$, where $N_{d,t}$ is the number treated areas in event year dataset t and N_d is the total number of treated areas across all event years. The three horizontal lines show weighted averages of the year-specific treatment effects using three sets of weights: (1) an unweighted average across event years; (2) a weighted average using the values shown on the x-axis; (3) a weighted average using the fraction of office closures (rather than treated areas) accounted for by each event year. Approach (2) is most similar, but not identical to, the baseline estimate in Table 4.

4.2 Caseload Composition

Understanding how closures affect caseload composition is important for normative interpretations of caseload declines, as discussed further in Section 5. In this section, I consider effect heterogeneity across demographics, education level, prior benefit history, immigrant status, and health status. I then estimate whether closures disproportionately discouraged applicants who would have had higher-than-average benefit durations.

To estimate effect heterogeneity, I subset the individual-level data according to some characteristic (*e.g.* immigrant status), construct the area-level outcome $\frac{y_{s,g,t}}{p_{s,g,t}}$ for each subset, and estimate the closure effect separately for each subset.

Figure 4 plots the static treatment effects on the combined IA and DA caseload for each subgroup. Figures C.3, C.4, and C.5 show the corresponding event study estimates. Table C.1 shows effects for IA and DA individually, along with p-values that formally test whether treatment effects are different between subgroups.

Heterogeneity by Age, Family Type, and Education Status: The results in Figure 4 and Table C.1 suggest a greater impact on young adults — an 11.7% decline among 18-25 year-olds versus a 7.1% decline among 26 to 40 year-olds. There is less evidence of a differential impact by family type (*i.e.*, childless adults versus parents). Even mothers with pre-school age children saw declines in IA use, although marginally less so than childless adults. This is notable because these mothers are exempt from work-search requirements[24] and presumably face larger barriers to labor force participation.

As described in Section 6, to examine heterogeneity by education, I focus on young adults, in part because the educational data begins in 1991, meaning educational attainment is poorly measured for older ages. Panel (a) of Figure 4 shows that caseload declines were strikingly similar between those with and without a high school (HS) certificate, *in percent terms*. This is another sign that office closures' effect were not strictly concentrated on less employable persons. With that said, the baseline rate of benefit receipt differs massively between the two groups: about 22% of young adults without a HS certificate received benefits in any given year, compared to only 3.15% among those who did graduate. So the effect of caseload declines *in levels* is far larger.

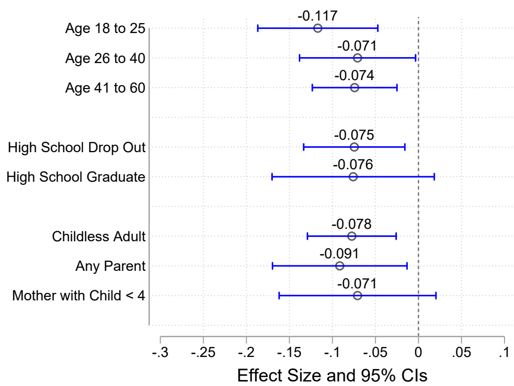
Heterogeneity by Proxies for Informational Frictions: To examine whether closures disproportionately affected persons who are less aware of the programs, or less able to navigate the application, I conduct three heterogeneity exercises, shown in panel (b) of Figure 4.

First, I proxy for awareness based on whether individuals have prior program exposure through their parents. Teenagers who are exposed to benefit programs via their parents are causally more

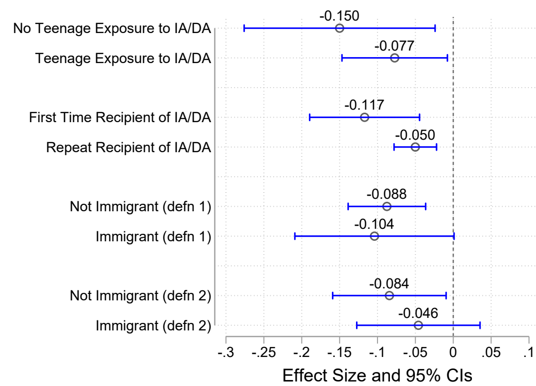
[24]They were also exempt from the two-week work-search during the application process starting around 2004.

Figure 4: **Heterogeneity in Combined IA & DA Declines**

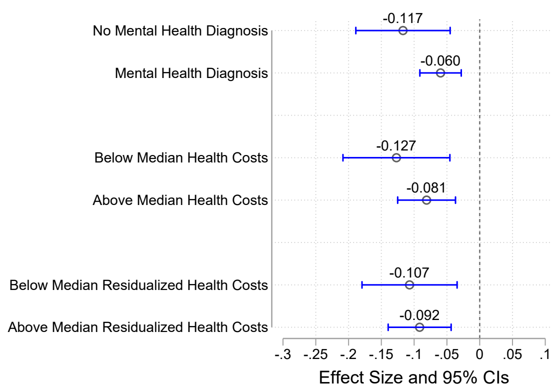
(a) Demographics



(b) Awareness



(c) Health



Note: Estimates of β from equation 1 for sub-sample splits and 95% confidence intervals. The dots are labelled with β. Each split is defined in the main text. Each panel groups the heterogeneity cuts into those related to demographics (panel (a)), those related to potential proxies for awareness (panel(b)), and those related to health status (panel (c)). Table C.1 shows the point estimates and p-values for the test of the equality of estimates in each split.

likely to use the programs as young adults (*e.g.*, De Haan and Schreiner (2025); Hicks et al. (2023); Hartley, Lamarche and Ziliak (2022); Dahl, Kostol and Mogstad (2014)), plausibly due to information transmission. Based on this logic, I test whether office closures disproportionately reduce young adult caseloads among those *without* prior family exposure compared to yonug adults with such exposure. I define teenage exposure as whether the young adult was in a household that received any IA or DA at any time from age 14 to age 17. This age range requires restricting the set of event closures, as outlined in the footnote.[25] I then estimate the effect on caseloads for young

[25] I exclude 18 and 19 because kids transition out of their parents' household at this age. Kids born in 1977 turn 14 in 1991, the first year of the caseload data, so 1977 is the first cohort for which I observe program receipt for the entire 14 to 17 age range. Because the 1977 cohort is age 24 in 2001, I only use office closures that occurred after 2001.

adults aged 20 to 24 who were born in 1977 or later.

The caseload decline for young adults without teenage exposure is 15% compared to only 7.7% for those with teenage exposure. This differential is consistent with the hypothesis that office closures disproportionally reduce caseloads among people with less program familiarity. However, the p-value on the difference is 0.234 (Table C.1), so this is at most suggestive. Also, young adults with teenage exposure differ in other ways that could mediate effect size – for example, they come from lower socio-economic backgrounds and may have less attractive employment or education opportunities as young adults. But, the heterogeneity cuts by high-school graduation status (panel (a)) suggest no effect size difference between young adults with, versus without, a high school certificate. In that light, the information friction interpretation is plausible.

Next, I examine effects among all working age adults based on whether they themselves received benefits at any time before the year of observation. The logic is the same: prior receipt signals greater awareness. Here again, caseload declines are larger for people without prior benefit history (in percent terms): an 11.7% drop in first-time recipients versus a 5% drop among prior recipients, which are statistically different at the 5% level (Table C.1). This adds more evidence to the information frictions interpretation, but again with the caveat that the two groups differ in other ways. This heterogeneity cut also relates to entries versus exits: people who received benefits last year are classified as prior recipients, and so examining whether they do not receive benefits in the current year captures program exits.

Finally, I explore heterogeneity by immigrant status. Immigrants may struggle with the application process (e.g., due to language barriers) and conceivably may be less aware of the programs. Both conjectures would predict that immigrants are disproportionately affected by closures. However, I do not find compelling effect heterogeneity between the two groups.

Heterogeneity by Health Status: Because health problems are a common barrier to work, health status is one proxy for a person's employability. The first measure I use is whether the person received medical care associated with a mental health diagnosis during the prior year. Mental health diagnoses are particularly common among IA and DA recipients (Green et al., 2021; Hicks, 2023) and represent plausibly work-limiting conditions in some cases. My second measure is the total healthcare costs incurred by the individual in the public health system in the prior year. Using total costs avoids taking a stance on which health issues are most relevant. I use costs in two ways. First, I split the sample into above- and below-median costs in the given year. Second, to isolate healthiness for a given life stage, I regress costs on fixed effects for age, gender, year, and their interactions, then split the sample into above- and below-median residualized costs.

Using contemporary health outcomes to split the sample may cause identification issues if closures affect health itself. In theory, this is testable by putting health measures as the outcome variable. But the exercise would be statistically underpowered since, as shown in Table 3, a closure lowers the percent of an area's population receiving benefits by only 0.45 percentage points. Usefully, there is evidence on this question: Hicks et al. (2023) and Hicks (2023) both find quite limited effects of benefit access on aggregate health costs, albeit with some effects on mental health. So while the identification challenge should not be ruled out, it may not be first-order. And to partly assuage it, I use one-year-lagged health outcomes, rather than current-year, to split the sample.

The estimates in panel (c) of Figure 4 show the results. Caseloads decline by 11.7% for people *without* a mental health diagnosis compared to only 6% for those with a mental health diagnosis. The p value on the difference is 0.04 (Table C.1). To the extent that mental health diagnoses represent work-limiting conditions, this finding is consistent with closures disproportionately lowering caseloads among more employable people. The finding also goes against the hypothesis that those with mental health diagnoses were disproportionately pushed off due to a lesser ability to navigate the application process without a nearby office.

A similar pattern exists when measuring health status by above vs below median health costs — those with lower healthcare costs saw larger percentage declines in benefit receipt (p-value on the difference is 0.063). However, this heterogeneity shrinks considerably when instead splitting the sample by *residualized* health costs. That is, the effect heterogeneity across health cost groups is primarily driven by age and gender-related effect heterogeneity.

A Measure of Changes in Targeting: A common measure of targeting, used by Finkelstein and Notowidigdo (2019) and Castell et al. (2024), is whether the complier households have different levels of benefit receipt, counterfactually, than do households who take up benefits despite an office closure. For example, both Castell et al. (2024) and Finkelstein and Notowidigdo (2019) find that application assistance disproportionately encourages enrollment among persons who receive lower benefit amounts once enrolled.

To adopt this test in my identification setup, I estimate equation 1, where the outcome is benefit duration and the sample is restricted to benefit recipients. So the estimated β reflects the change in average benefit duration among recipients in treated areas relative to control areas. Assuming closures only affect the extensive margin of receipt, not the intensive margin, the change in average benefit durations reflects changes in caseload composition. For example, if closures discourage applicants who would have had above-average benefit duration, then β will be negative.

Table 5: **Closures' Effect on Annual Benefit Amounts Among Recipients**

	Months Received		Dollars Received	
	IA	DA	IA	DA
Treated $\times$ Post	-0.00437	0.00246	-0.00345	-0.00620
	(0.00705)	(0.00269)	(0.0103)	(0.00490)
Observations	2938	2923	2938	2923

Note: This table contains estimates of β from the model $E\left(y_{s,g,t}\right) = \exp\left(\alpha_{g,t} + \alpha_{s,t} + \beta\mathbb{1}\{s > 0\}D_{g,t}\right)$ where y is average annual benefits received (in months or dollars) among benefit recipients. *** $\rho < .01$, ** $\rho < .05$, * $\rho < .1$.

The results are shown in Table 5. The first two columns measure benefit duration by the number of months received during the year. The results are precisely-null, suggesting zero change in targeting along this dimension. The second two columns use annual dollar amounts instead, and again, closures have precisely zero effect. There is no evidence of significant changes in targeting.

5 Discussion

A large literature has documented that benefit take-up is inversely proportional to the degree of complexity, hassle, and imperfect awareness (see Herd et al. (2023) for reviews). On the high end, the Earned Income Tax Credit has a take-up rate of 80% (Goldin and Liscow, 2018). In contrast, welfare and disability assistance often have much lower take-up[26], presumably due to their greater complexity and hassles. Given this complexity, it is important to understand how application modes, such as field offices, affect take-up and targeting. My paper contributes to this question.

Economic theory suggests that application ordeals can improve aggregate welfare if they disproportionately screen-out “high ability" individuals (Nichols and Zeckhauser, 1982; Besley and Coate, 1992; Diamond and Sheshinski, 1995) and, therefore, the empirical literature has emphasized the targeting properties of related interventions.[27] For example, both Castell et al. (2024) and Finkelstein and Notowidigdo (2019) find that application assistance disproportionately increases take-up among less eligible households[28], while in contrast, Deshpande and Li (2019) find that office closures worsen targeting efficiency in disability assistance. In comparison to those studies, my results are in between: within welfare, caseload declines were widespread across age, health status, education levels, immigrant status, and eligibility groups — albeit there is suggestive evidence of modestly larger declines among younger and healthier people. Closures also did not impact the caseload composition as measured by benefit duration. And I find generally limited declines in disability assistance caseloads.

[26] Crouse and Waters (2014) use survey data to estimate that TANF take-up rates range from 51.8% in 2000 to 32.8% in 2012 but Meyer and Mittag (2019) demonstrate that errors in survey data significantly bias downward inferred take-up. Hernanz, Malherbet and Pellizzari (2004) and Ko and Moffitt (2024) survey take-up rates for similar programs in other countries. Whelan (2010)'s estimates of take-up for Canadian social assistance range from 13% to 47%.

[27] Experimental and quasi-experimental work has aimed to separate the roles of imperfect information versus application hassles (Bettinger et al., 2012; Bhargava and Manoli, 2015; Armour, 2018; Finkelstein and Notowidigdo, 2019; Castell et al., 2024), sometimes finding both to be important, depending on the study. Alatas et al. (2016) find that just requiring an application for benefits (as opposed to automatic enrollment) improved the targeting efficiency of an Indonesian program, but slightly increasing the costliness of the application did not further affect targeting.

[28] In Finkelstein and Notowidigdo (2019), this means higher-income and healthier individuals had the larger boosts to take-up of SNAP following information and assistance interventions. In Castell et al. (2024), this means that take-up increased more for persons eligible for only small dollar amounts of benefits.

An important interpretation point is that my identification strategy takes as given civil society responses. When a government office closes, and therefore reduces in-community support, local non-profits, political constituency staff, or even public librarians may step-up support to potential applicants. Anecdotal evidence is consistent with this compensatory behavior. So my estimated effects are a lower bound on the "all-else-equal" effect of an office closure, but are arguably more policy-relevant. Civil society responses may also affect the targeting properties of an office closure — for example, by disproportionately assisting the most vulnerable individuals. This hypothesis could explain why office closures did not worsen proportional targeting.[29]

References

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Ririn Purnamasari, and Matthew Wai-Poi. 2016. "Self-Targeting: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy*, 124(2): 371–427.

Anders, Jenna, and Charlie Rafkin. 2024. "The Welfare Effects of Eligibility Expansions: Theory and Evidence from SNAP." *American Economic Journal: Economic Policy*.

Armour, Philip. 2018. "The Role of Information in Disability Insurance Application: An Analysis of the Social Security Statement Phase-In." *American Economic Journal: Economic Policy*, 10(3): 1–41.

Besley, Timothy, and Stephen Coate. 1992. "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs." *The American Economic Review*, 82(1): 249–261.

Bettinger, Eric P, Bridget Terry Long, Philip Oreopolous, and Lisa Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment." *The Quarterely Journal of Economics*, 1205–1242.

Bhargava, Saurabh, and Dayanand Manoli. 2015. "Psychological Frictions and the Incomplete Take-Up of Social Benefits : Evidence from an IRS Field Experiment." *American Economic Review*, 105(11): 3489–3529.

Biscoe, Chris, Christina McMillan, Wendy Byrne, Zoe Macmillian, Kate Morrison, and Lisa Phillips. 2018. "Holding Pattern: Call Wait Times for Income and Disability Assistance." The

[29] Another interpretation point is that reduced take-up caused by closures could capture both a "direct effect" where people don't take up without in-person support and an "indirect" effect where reduced caseloads reduce awareness in the community, further reducing caseloads. Anders and Rafkin (2024) find evidence of an "indirect" effect in SNAP.

Office of the Ombudsperson.

Blackwell, Matthew, Stefano Iacus, Gary King, and Giuseppe Porro. 2009. "cem: Coarsened exact matching in Stata." *Stata Journal*, 9(4): 524–546.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2024. "Difference-in-Differences with a Continuous Treatment."

Castell, Laura, Marc Gurgand, Clement Imbert, and Todor Tochev. 2024. "Take-up of Social Benefits: Experimental Evidence from France." *American Economic Journal: Economic Policy*.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. 2019. "The Effect of Minimum Wages on Low-Wage Jobs*." *The Quarterly Journal of Economics*, 134(3): 1405–1454.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin. 2020. "Fast Poisson estimation with high-dimensional fixed effects." *The Stata Journal*, 20(1): 95–115.

Crouse, Gilbert, and Annette Waters. 2014. "Welfare Indicators and Risk Factors." U.S. Department of Health and Human Services, Washington, D.C.

Dahl, Gordon B, Andreas Ravndal Kostol, and Magne Mogstad. 2014. "Family Welfare Cultures." *Quarterely Journal of Economics*, 129(4): 1711–1752.

De Haan, Monique, and Ragnhild C Schreiner. 2025. "The Intergenerational Transmission of Welfare Dependency." *The Economic Journal*, 135(669): 1609–1640.

Deshpande, Manasi, and Yue Li. 2019. "Who Is Screened Out? Application Costs and the Targeting of Disability Programs." *American Economic Journal: Economic Policy*, 11(4): 213–248.

Diamond, Peter, and Eytan Sheshinski. 1995. "Economic aspects of optimal disability benefits." *Journal of Public Economics*, 57: 1–23.

Finkelstein, Amy, and Matthew J Notowidigdo. 2019. "Take-up and targeting: experimental evidence from SNAP." *Quarterely Journal of Economics*, 1505–1556.

Foote, Andrew, and Stephanie Rennane. 2019. "The Effect of Lower Transaction Costs on Social Security Disability Insurance Application Rates and Participation." *Journal of Policy Analysis and Management*, 38(1): 99–123.

Frenette, Marc, Yuqian Lu, and Grant Schellenberg. 2015. "Social Assistance Receipt Among Refugee Claimants in Canada."

Giannella, Eric, Tatiana Homonoff, Gwen Rino, and Jason Somerville. 2023. "Administrative Burden and Procedural Denials: Experimental Evidence from SNAP." *American Economic Journal: Economic Policy*.

Goldin, Jacob, and Zachary Liscow. 2018. "Tax Benefit Complexity and Take-Up: Lessons from the Earned Income Tax Credit." *Tax Law Review*, 72(59).

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing."

Journal of Econometrics, 225(2): 254–277. Themed Issue: Treatment Effect 1.

Gourieroux, C., A. Monfort, and A. Trognon. 1984. "Pseudo Maximum Likelihood Methods: Applications to Poisson Models." *Econometrica*, 52(3): 701–720.

Green, David, Jeffrey Hicks, Rebecca Warburton, and William Warburton. 2021. "BC Income Assistance Trends and Dynamics: Descriptions and Policy Implications." *Report commissioned by the Expert Panel on Basic Income*.

Green, David, Rhys Kesselman, and Lindsay Tedds. 2021. "Covering All the Basics Reforms for a More Just Society: Final Report of the British Columbia Expert Panel on Basic Income."

Hartley, Robert Paul, Carlos Lamarche, and James P. Ziliak. 2022. "Welfare Reform and the Intergenerational Transmission of Dependence." *Journal of Political Economy*, 130(3): 523–565.

Herd, Pamela, Hilary Hoynes, Jamila Michener, and Donald Moynihan. 2023. "Introduction: Administrative Burden as a Mechanism of Inequality in Policy Implementation." *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 9(4): 1–30.

Hermes, Henning, Philipp Lergetporer, Frauke Peter, and Simon Wiederhold. 2024. "Application Barriers and the Socioeconomic Gap in Child Care Enrollment." *Journal of the European Economic Association*, 23(3): 1133–1172.

Hernanz, Virginia, Franck Malherbet, and Michele Pellizzari. 2004. "Take-Up of Welfare Benefits in OECD Countries." , (17).

Hicks, Jeffrey. 2023. "Cash Welfare and Health Spending." *SSRN Electronic Journal*.

Hicks, Jeffrey, Gaelle Simard-Duplain, David Green, and William Warburton. 2023. "The Effect of Reducing Welfare Access on Employment, Health, and Children's Long-Run Outcomes." *SSRN Electronic Journal*.

Homonoff, Tatiana, and Jason Somerville. 2021. "Program Recertification Costs: Evidence from SNAP." *American Economic Journal: Economic Policy*, 13(4): 271–98.

Klein, Seth, Marge Reitsma-Street, and Bruce Wallace. 2006. "Denied Assistance: Closing the Front Door on Welfare in BC." Canadian Center for Policy Alterantives March.

Kleven, Henrik Jacobsen, and Wojciech Kopczuk. 2011. "Transfer Program Complexity and the Take-Up of Social Benefits." *American Economic Journal: Economic Policy*, 3(1): 54–90.

Kneebone, Ronald, and Katherine White. 2014. "The Rise and Fall of Social Assistance-Use in Canada, 1969-2012." *SPP Research Papers*.

Ko, Wonsik, and Robert A. Moffitt. 2024. "Take-Up of Social Benefits." *Handbook of Labor, Human Resources and Population Economics*, , ed. Klaus F. Zimmermann, 1–42. Cham:Springer International Publishing.

Meyer, Bruce D., and Nikolas Mittag. 2019. "Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness, and Holes in the Safety Net." *American Economic Journal: Applied Economics*, 11(2): 176–204.

Nichols, Albert L, and Richard J Zeckhauser. 1982. "Targeting Transfers through Restrictions on Recipients." *American Economic Review: Papers and Proceedings*, 72(2): 372–377.

Peterson, Sandra, Maeve Wickham, Ruth Lavergne, Jonathan Beaumier, Megan Ahuja, Dawn Mooney, and Kimberlyn McGrail. 2021. "Methods to comprehensively identify emergency department visits using administrative data in British Columbia." UBC Centre for Health Services and Policy Research.

Rafkin, Charlie, Adam Solomon, and Evan J. Soltas. 2023. "Self-Targeting in U.S. Transfer Programs." SSRN Working Paper 4495537, Cambridge, MA. Last revised April 30, 2024; 100 pp.

Ronayne, Bruce, Cary Chiu, Carly Hyman, Alyne Mochan, Sarah Barnes, David Gagnon, Zoe Macmillan, and Karen Sawatzky. 2009. "Last Resort: Improving Fairness and Accountability in British Columbia's Income Assistance Program." British Columbia: Office of the Ombudsman 45, Victoria, British Columbia.

Rossin-Slater, Maya. 2013. "WIC in your neighborhood: New evidence on the impacts of geographic access to clinics." *Journal of Public Economics*, 102: 51–69.

Roth, Jonathan, and Jiafeng Chen. 2023. "Logs with zeros? Some problems and solutions."

Roth, Jonathan, Pedro H.C. Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's trending in difference-in-differences? A synthesis of the recent econometrics literature." *Journal of Econometrics*, 235(2): 2218–2244.

Silva, J.M.C. Santos, and Rainer Winkelmann. 2024. "Misspecified Exponential Regressions: Estimation, Interpretation, and Average Marginal Effects." *The Review of Economics and Statistics*, 1–25.

Silva, J. M. C. Santos, and Silvana Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and Statistics*, 88(4): 641–658.

Stepner, Michael. 2013. "BINSCATTER: Stata module to generate binned scatterplots."

Whelan, Stephen. 2010. *Empirical Economics*, 847–875.

Wing, Coady, Seth M Freedman, and Alex Hollingsworth. 2024. "Stacked Difference-in-Differences." National Bureau of Economic Research Working Paper 32054.

Wooldridge, Jeffrey M. 2023. "Simple approaches to nonlinear difference-in-differences with panel data." *The Econometrics Journal*, 26(3): C31–C66.

Wu, Derek, and Bruce D Meyer. 2023. "Certification and Recertification in Welfare Programs:

What Happens When Automation Goes Wrong?" National Bureau of Economic Research Working Paper 31437.

Zuo, George, and David Powell. 2023. "The Impact of Affordable and Accessible Broadband on SSDI and SSI Participation." *University of Michigan Retirement and Disability Research Center (MRDRC) Working Paper; MRDRC WP 2023-467*.

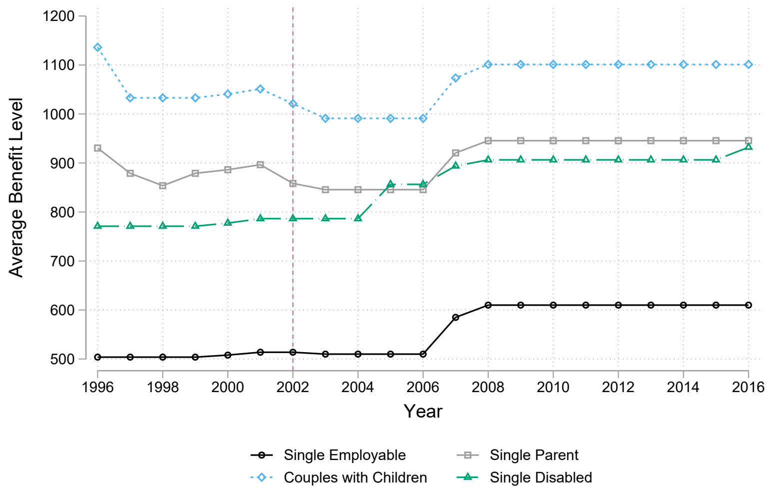
6 Data Citations

- British Columbia Ministry of Social Development and Poverty Reduction [creator] (2021): BC Employment and Assistance (BCEA) V04. Data Innovation Program, Province of British Columbia [publisher]. Data Extract. Approver Year 2023.
- British Columbia Ministry of Health [creator] (2021): Registration and Premium Billing (RPBLite) V02. Data Innovation Program, Province of British Columbia [publisher] Data Extract. Approver Year 2023.
- British Columbia Ministry of Health [creator] (2021): Discharge Abstract Database (DAD). Data Innovation Program, Province of British Columbia [publisher] Data Extract. Approver Year 2023.
- British Columbia Ministry of Health [creator] (2021): MSP Payment Information. Data Innovation Program, Province of British Columbia [publisher] Data Extract. Approver Year 2023.
- British Columbia Ministry of Health [creator] (2021): Consolidation File. Data Innovation Program, Province of British Columbia [publisher] Data Extract. Approver Year 2023.
- British Columbia Ministry of Health [creator] (2021): Vital Statistics Births. Data Innovation Program, Province of British Columbia [publisher] Data Extract. Approver Year 2023.
- British Columbia Ministry of Education and Child Care [creator] (2021): K to 12 Student Demographics and Achievements. Data Innovation Program, Province of British Columbia [publisher] Data Extract. Approver Year 2023.

APPENDIX FOR ONLINE PUBLICATION

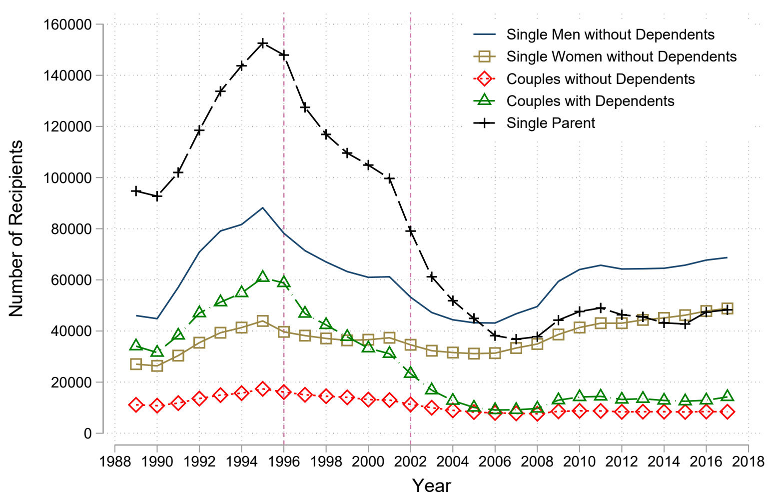
A Institutional Details

Figure A.1: **Nominal Monthly Benefit Rates By Year and Group**



Note: Nominal average monthly benefit rates by benefit type and group. The first three groups reflect welfare rates. Only "Single Disabled" represents disability assistance rates, in this case for a single-headed household which is the most common type on disability assistance.

Figure A.2: **Combined IA and DA Caseloads in British Columbia by Household Type**



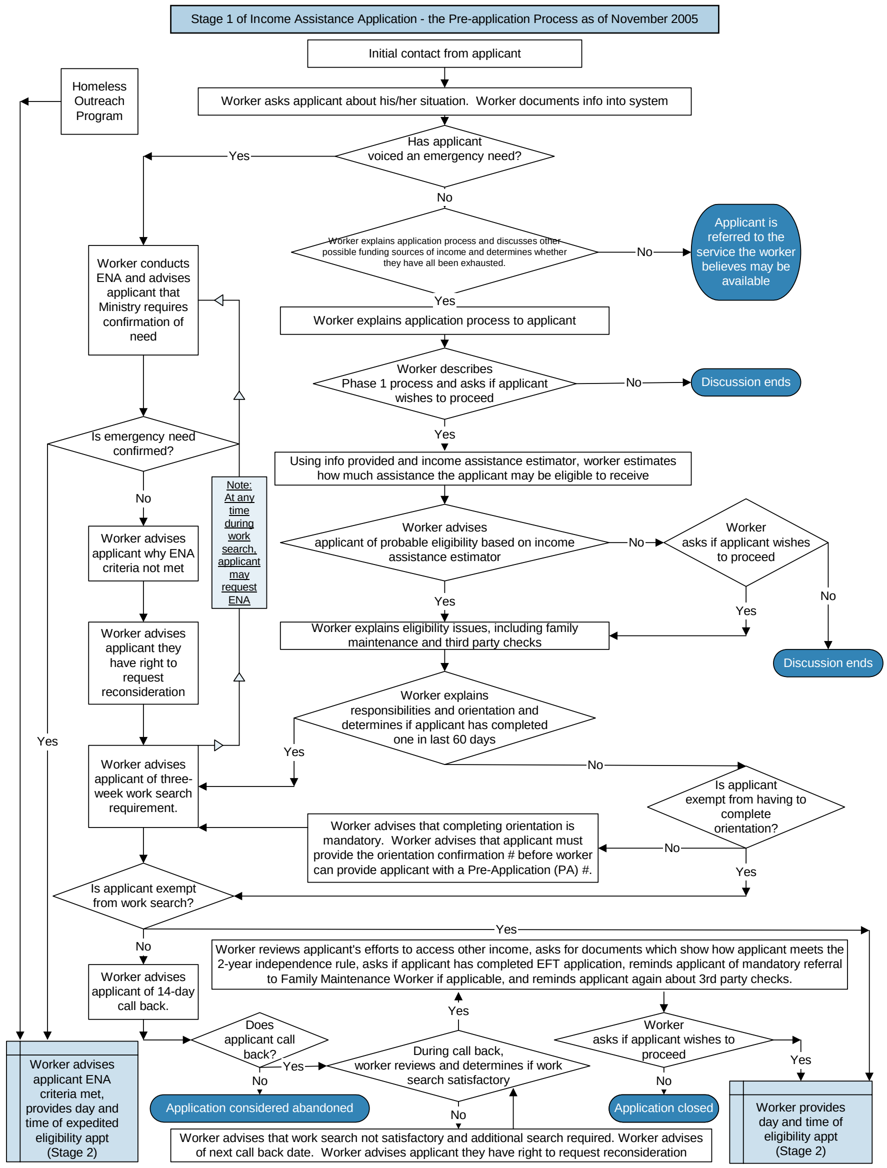
Note: This figure shows the number of persons receiving IA or DA as of July 1st in the given year. Persons are split according to household type. The vertical lines denote two periods of discrete policy changes that were intended to reduce access to these programs.

Figure A.3: **Application Process: 2002-2007**

OFFICE OF THE OMBUDSMAN 103

Appendix B

Income Assistance Application Process before October 29, 2007



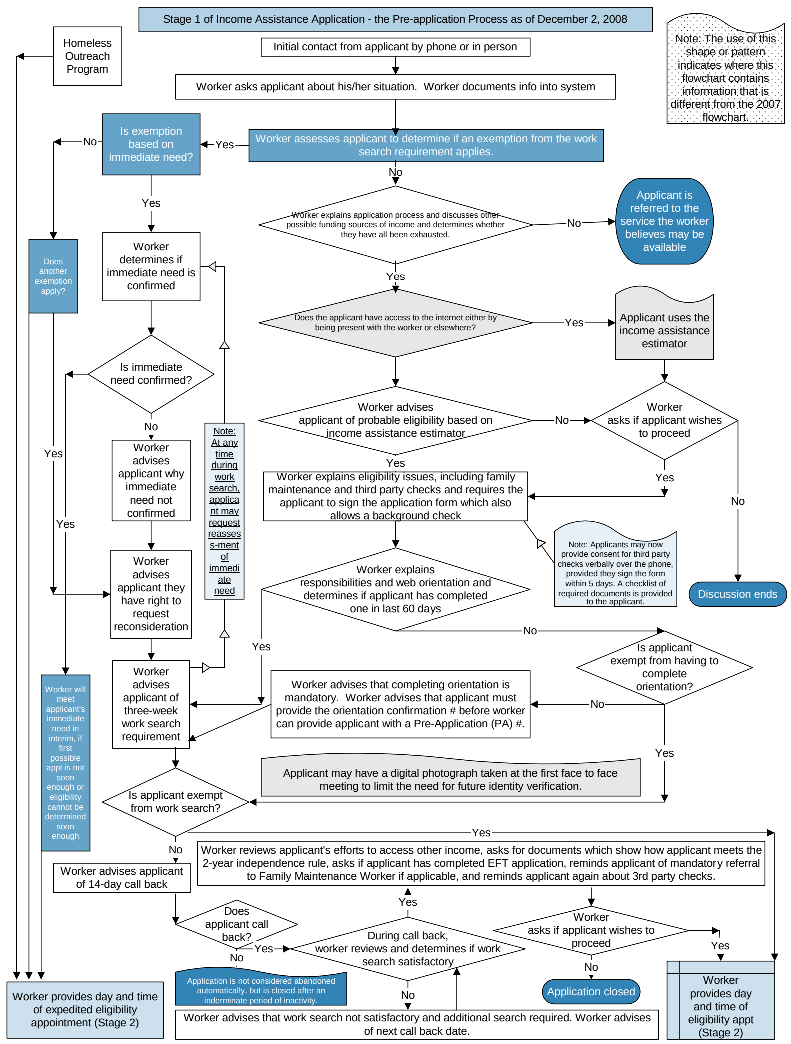
Source: (Ronayne et al., 2009)

Figure A.4: **Application Process: 2008 -Present**

OFFICE OF THE OMBUDSMAN 105

Appendix D

Income Assistance Application Process after December 2, 2008



Source: (Ronayne et al., 2009)

B Data Appendix

B.1 Office Locations and Closures

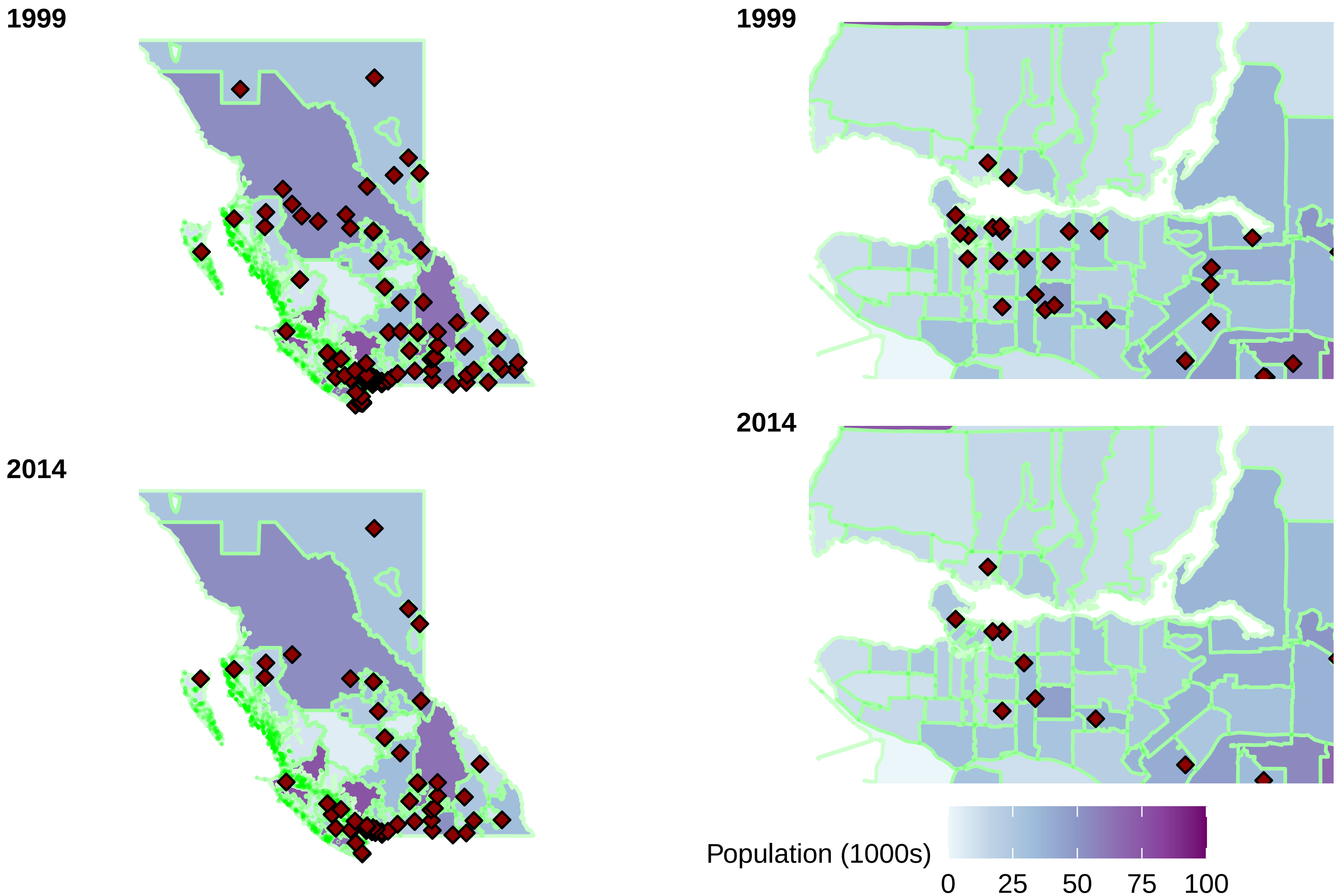
For the years 1991 through 1994, I collect office locations from **BC Guide**, a directory of government services published by Enquiry BC, and now accessible at the Royal BC Archives. For 1995, 1997, 1998, and 2000 I collect the locations from **Fast Track: A Guide to the Ministry**. This publication was an annually updated resource produced by the ministry responsible for IA and DA and distributed to MLA (provincial elected representatives) offices to provide them with up-to-date information they could use to direct constituents to the appropriate resources. Both the **BC Guide** and **Fast Track** provide information as of roughly June in the calendar year. For 1999, and 2001 to 2019, I use archived web pages from the ministry which published office locations online. The archives are stored by the Wayback Machine accessible at `https://archive.org/web/`. To be consistent with the physical sources, I use archived web-pages from summer months whenever possible. A list of the URLs and Python code for scraping will be published in the code repository.

I collect a list of publicized closures from 2002 to 2004 from web archived announcements available here: `https://web.archive.org/web/20031221213318/http://www.mhr.gov.bc.ca/officechanges/summary_reg.htm`. I complement this list with two other lists of closures post-2004 available in Ombudsperson reports and complaints. Biscoe et al. (2018) lists six closures between 2013 and 2014 on page 11. A report by the B.C. Public Interest Advocacy Center (BCPIAC) to the Ombudsperson of British Columbia lists another set of closures between 2005 and 2013, available here: `https://web.archive.org/web/20220430001944/https://bcpiac.com/wp-content/uploads/2015/09/BCPIAC-Ombuds-Complaint_Final_May-12-2015.pdf`. Both of these sources are cross-checked against the public listings of offices in operation to ensure accuracy. Because most closures were simply not announced, I then infer all other closures by comparing year-over-year changes in the list of offices in operation. A final master list of closures used for this project will be available in the replication package.

B.2 Defining Areas and Calculating Distances

Postal Code Conversion Files (PCCF) link all six-digit postal codes (PC) to the latitude and longitude corresponding to the PC centroid. I download PCCF files corresponding to 1991, 1996, 2001, 2006, 2011, and 2016 to obtain a panel of PC centroids. I match each centroid to the nearest office latitude and longitude using straight-line distance.

Figure B.1: **Office Locations**



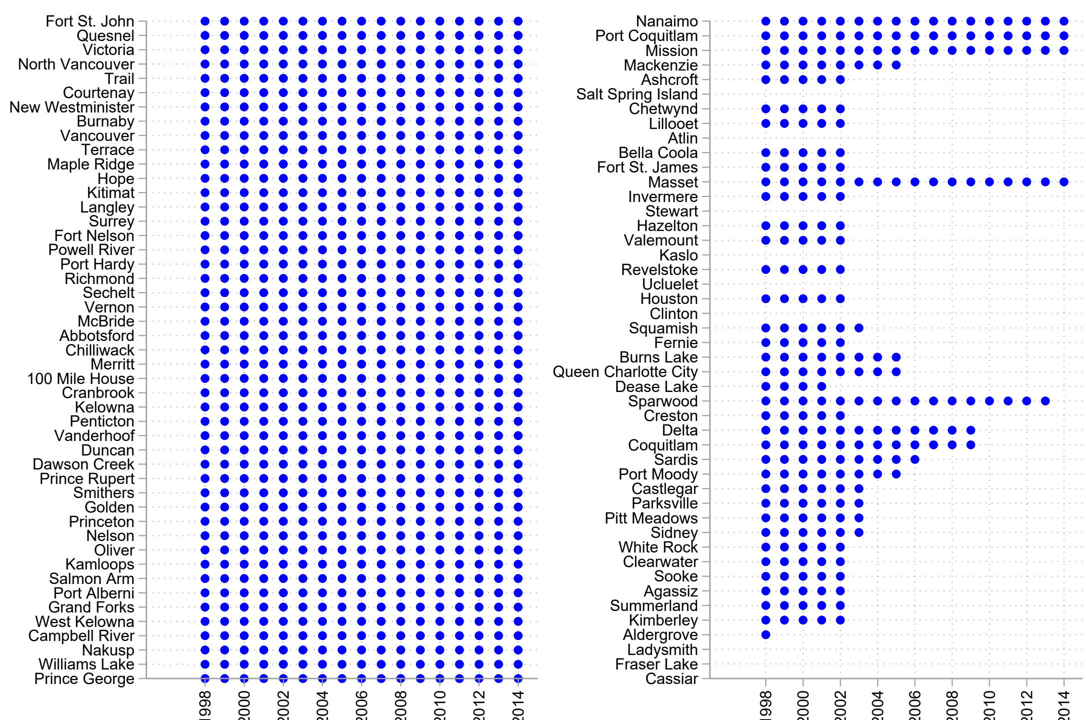
Note: The two left panels map the entire province of British Columbia and the two right panels depict the Greater Vancouver Area. The green lines illustrate the boundaries of FSAs as of 2011. The shading illustrates the distribution of population across FSAs as of 2011 (population by FSA is not available for prior years). Rural FSAs are geographically larger which offsets the lower population density. The red diamonds indicate field offices in operation; the top panels show the availability of offices in 1999, and the bottom two panels as of 2014.

Table B.1: **Counts for Larger Communities**

	1999	2014
Vancouver	21	7
Victoria	7	2
Surrey	6	4
Nanaimo	4	2
Kamloops	4	3
Prince George	4	1
Burnaby	4	1
Abbotsford	3	1
Kelowna	3	1
Richmond	3	1
North Vancouver	3	1

Note: This table shows the number of offices, defined by their administrative code rather than by unique location, for communities that had at least two offices at some point in time. Appendix B describes the original sources.

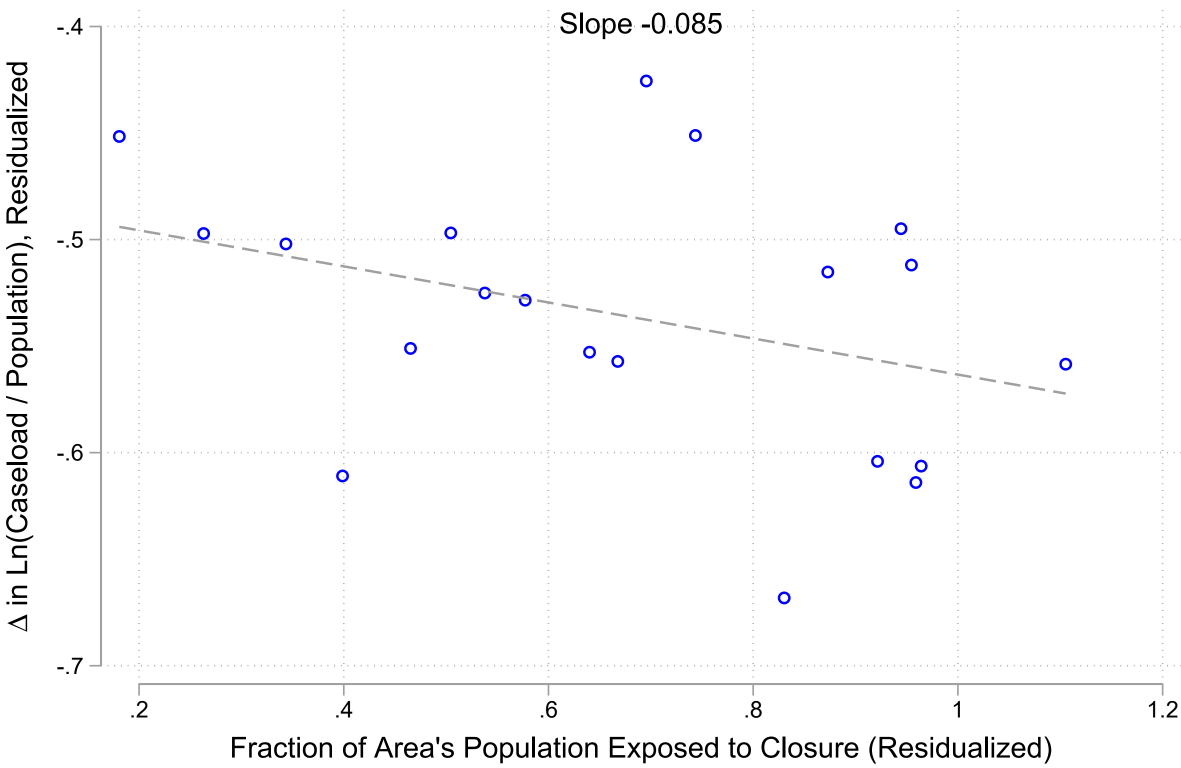
Figure B.2: **Presence of At Least One Office by Community and Year**



Note: For each community and year, this figure indicates whether at least one office was in operation. Office closures occurred in rural and urban areas, but most urban areas have many offices and so show a consistent presence throughout the period. Alternatively, closures in rural communities resulted in those communities having no access to in-person services.

C Supplemental Results

Figure C.1: **Change in Caseloads Against the Fraction of the Area Exposed to the Closure**



Note: The horizontal axis is the fraction of the treated area's population exposed to the closure (measurement of this is described in Section 2). The vertical axis is the change in the area's combined IA and DA caseload from before to after the closure, measured as $\sum_{s=1}^{4} ln(\frac{y_{s,g,t}}{p_{s,g,t}}) - \sum_{s=-4}^{-2} ln(\frac{y_{s,g,t}}{p_{s,g,t}})$. The dots show the binscatter (Stepner, 2013) relationship between the two after controlling for the year of closure. The downward-sloping fitted line indicates that areas with greater exposure to the closure saw larger caseload declines. The slope of the fitted line is γ in the following regression: $\frac{\sum_{s=1}^{4} ln(\frac{y_{s,g,t}}{p_{s,g,t}})}{4} - \frac{\sum_{s=-4}^{-2} ln(\frac{y_{s,g,t}}{p_{s,g,t}})}{3} = \gamma \text{Fraction of Population Exposed}_{g,t} + \gamma_t + \epsilon_{g,t}$, where γ_t are fixed effects for closure year.

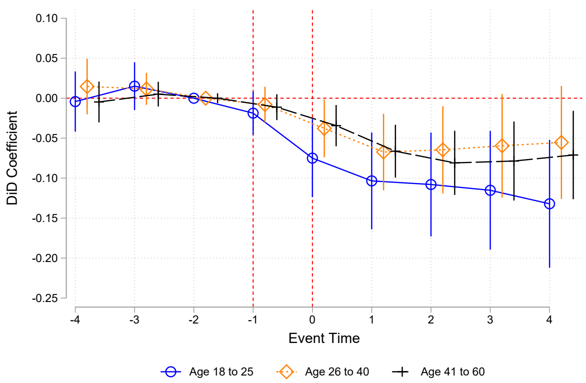
Table C.1: **Effects by Subgroup**

	Any IA/DA			IA			DA		
Subgroup	β	SE	ρ	β	SE	ρ	β	SE	ρ
Demographics									
Age 18 to 25	-0.117	0.036		-0.159	0.046		-0.056	0.032	
Age 26 to 40	-0.071	0.034		-0.088	0.046		0.005	0.035	
Age 41 to 60	-0.074	0.025		-0.112	0.042		0.004	0.027	
Childless Adult	-0.078	0.026		-0.115	0.043		-0.018	0.023	
Any Parent	-0.092	0.040		-0.098	0.046		-0.049	0.036	
Mother with Child < 4	-0.071	0.047		-0.077	0.049		-0.030	0.070	
Did Not Receive High School	-0.075	0.030		-0.125	0.041		-0.004	0.0415	
Received HS Diploma	-0.076	0.048	0.974	-0.084	0.059	0.385	-0.021	0.071	0.823
Proxies for Information Frictions									
No Teenage Exposure to IA/DA	-0.150	0.064		-0.191	0.078		0.030	0.080	
Teenage Exposure to IA/DA	-0.077	0.036	0.234	-0.120	0.046	0.265	0.013	0.049	0.862
Repeat Recipient of IA/DA	-0.050	0.014		-0.078	0.028		0.007	0.016	
First Time Recipient of IA/DA	-0.117	0.037	0.025	-0.139	0.043	0.021	0.003	0.044	0.908
Not Immigrant (defn 1)	-0.088	0.026		-0.115	0.041		-0.007	0.029	
Immigrant (defn 1)	-0.104	0.054	0.693	-0.095	0.058	0.646	-0.076	0.066	0.273
Not Immigrant (defn 2)	-0.084	0.038		-0.096	0.043		0.014	0.038	
Immigrant (defn 2)	-0.046	0.042	0.365	-0.038	0.042	0.205	-0.104	0.095	0.215
Health Status									
Below Median Health Costs	-0.127	0.042		-0.154	0.054		-0.030	0.034	
Above Median Health Costs	-0.081	0.023	0.063	-0.102	0.036	0.044	-0.007	0.027	0.238
< Median Residualized Health Costs	-0.107	0.037		-0.135	0.051		-0.020	0.032	
≥ Median Residualized Health Costs	-0.092	0.025	0.436	-0.111	0.038	0.253	-0.017	0.026	0.918
No Mental Health Diagnosis	-0.117	0.037		-0.149	0.053		-0.040	0.031	
Mental Health Diagnosis	-0.060	0.016	0.040	-0.083	0.028	0.041	0.014	0.024	0.006

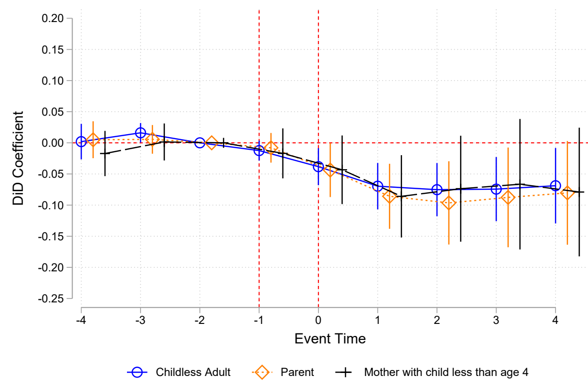
Note: This table shows estimates of β from equation 1 for different subgroups of the data, as described in Section 4.2. ρ tests the null hypothesis that β is identical in each pairwise subgroup. The first three columns correspond to the combined IA and DA caseload. The second and third sets of columns correspond to effects on IA and DA caseloads individually.

Figure C.2: **Heterogeneity by Demographics**

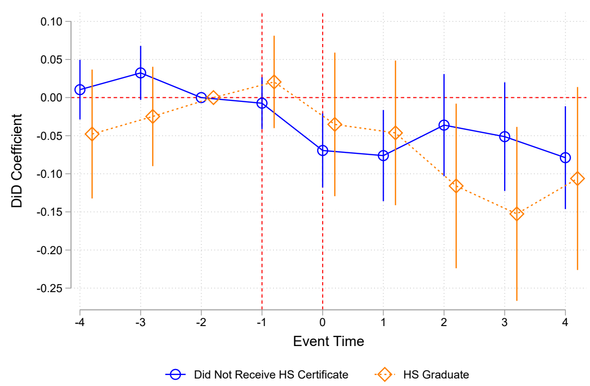
(a) **By Age Group**



(b) **Childless Adults vs Parents**



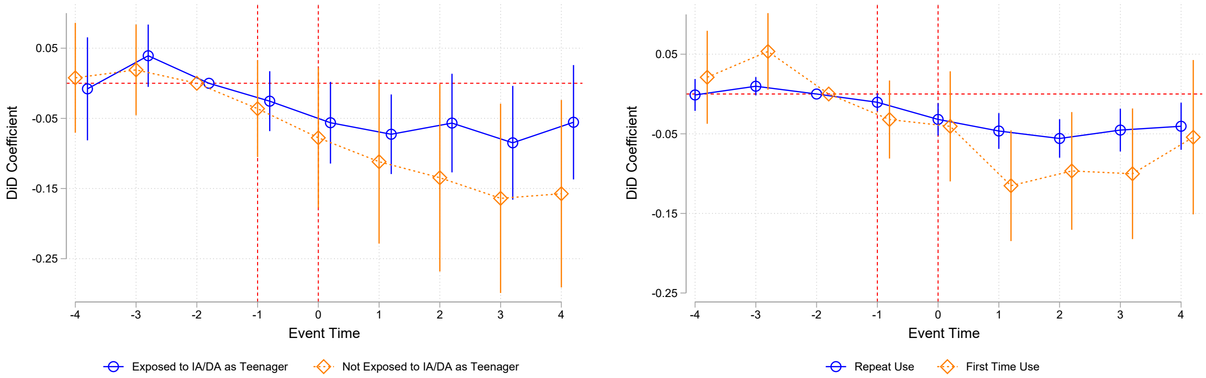
(c) **By High School (HS) Graduation**



Note: This figure plots estimates of β_s from equation 3 for each s from -4 to 4, and 95% confidence intervals, for the combined IA & DA caseload. Each area's population is first split into groups based on observable characteristics, then equation 3 is estimated on each sub-sample. The red vertical lines denote the office closure time between -1 and 0. Only results from the matched control group are shown. Panel (a) splits the sample into three age groups. Panel (b) splits the sample into childless adults vs parents, in addition to showing results specifically for mothers with children less than age 4. Panel (c) splits the sample according to high-school graduation status among the sample of young adults (age 20 - 24) who were observed in the enrollment records as teenagers.

Figure C.3: **Heterogeneity by Previous Program Exposure**

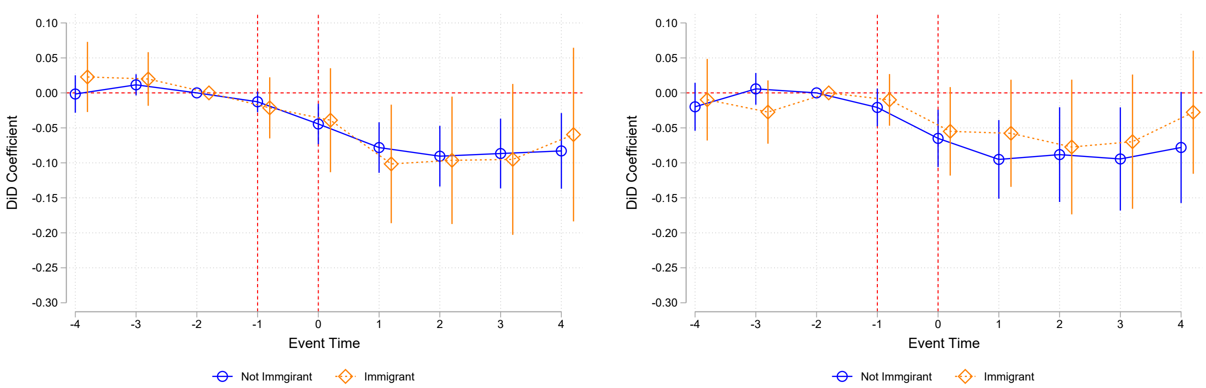
(a) **By Teenage Exposure** (b) **By Previous Personal Receipt**



Note: This figure plots estimates of β_s from equation 3 for each s from -4 to 4, and 95% confidence intervals, for the combined IA & DA caseload. Standard errors are clustered at the area level. Each area's population is first split into groups based on observable characteristics, then equation 3 is estimated on each sub-sample. The red vertical lines denote the time of office closure, which happens sometime between -1 and 0. Panel (a) splits the sample by whether the person was exposed to IA/DA during their teenage years (age 14 to 17), as described in Section 4.2. Panel (b) uses the full adult sample, split by whether the adult had prior benefit receipt.

Figure C.4: **Heterogeneity by Immigrant Status**

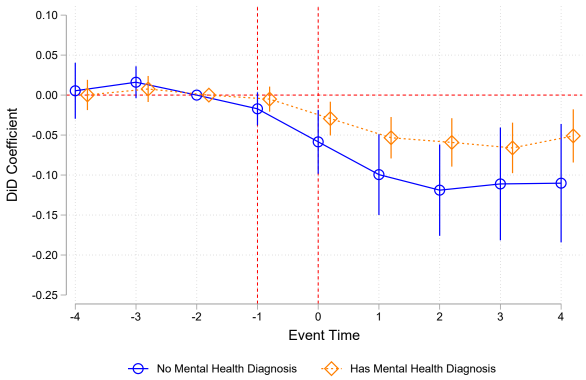
(a) **Immigrant Definition 1** (b) **Immigrant Definition 2**



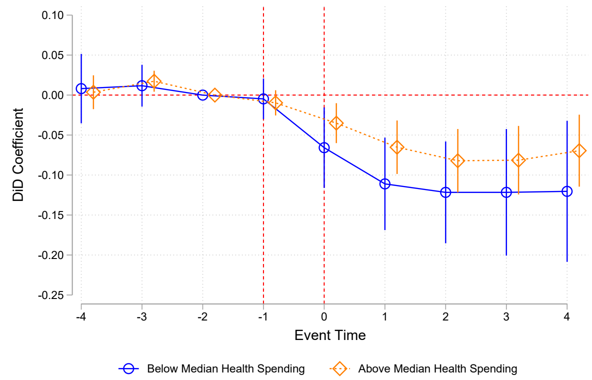
Note: This figure plots estimates of β_s from equation 3 for each s from -4 to 4, and 95% confidence intervals, for the combined IA & DA caseload. Standard errors are clustered at the area level. Each area's population is first split into groups based on observable characteristics, then equation 3 is estimated on each sub-sample. The red vertical lines denote the time of office closure, which happens sometime between -1 and 0. Panel (a) splits the sample according to the first definition of immigrant status. Panel (b) splits the sample according to the second definition of immigrant status. Both definitions are described in Section 2.

Figure C.5: **Heterogeneity by Health Status**

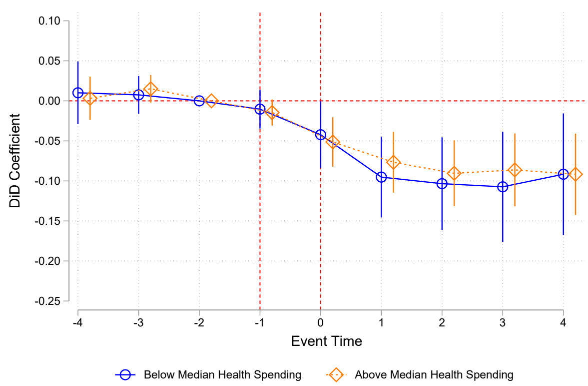
(a) **By Mental Health Diagnosis**



(b) **By Healthcare Costs**



(c) **By Residualized Healthcare Costs**



Note: This figure plots estimated β_s from equation 3 for each s from -4 to 4, and 95% confidence intervals, for the combined IA & DA caseload. Each area's population is first split into groups based on observable characteristics, then equation 3 is estimated on each sub-sample. The red vertical lines denote the time of office closure between -1 and 0. Only results from the matched control group are shown. Panel (a) splits by the presence of a mental health diagnosis in the year. Panel (b) divides persons into above vs. below median healthcare costs. Panel (c) divides persons into above vs. below median *residualized* healthcare costs.