Essays on Public Policy and Economic Inequality

AN ABSTRACT

SUBMITTED ON THE EIGHTEENTH DAY OF APRIL, 2017

TO THE DEPARTMENT OF ECONOMICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

OF THE SCHOOL OF LIBERAL ARTS

OF TULANE UNIVERSITY

FOR THE DEGREE

OF

DOCTOR OF PHILOSOPHY

BY

Daniel Teles

APPROVED: Alan Barreca, Ph.D.
CHAIR

Stefano Barbieri, Ph.D.

Steven Sheffrin, Ph.D.

Abstract

This dissertation consists of three essays. The first is an analysis of tax credits that encourage charitable giving that estimates the causal effect of two of the costliest program: Arizona's Working Poor Tax Credit (WPTC) and the Endow Iowa program. Using synthetic control methods to construct counterfactuals, I estimate a 125 percent increase in contributions to community foundations in Iowa. In contrast, I find no evidence that the WPTC increased contributions to the targeted Arizona nonprofits. Evidence suggests that the growth in contribution levels in Iowa included increases in both the number of foundations and the level of contributions per foundation.

The second essay examines the economic impact of the AmeriCorps program on the nonprofits. Fixed effects regressions show that nonprofits experience higher levels of contributions in years in which they sponsor AmeriCorps, whether State and National or VISTA. However, instrumental variables analysis suggests that AmeriCorps State and National have a crowding out effect on contributions. A ten percent increase in the number of AmeriCorps State and National is associated with a one percent decline in contributions. This level of crowd-out is similar to those estimated for other forms of government funding for the nonprofit sector. Estimates of the impact of VISTA rule out large levels of crowd-out.

The third essay tests the sensitivity of cross-national inequality research to the choices about the underlying data. The main takeaways are as follows. First, estimates appear to be more sensitive to the choice of welfare concept than to the choice of inequality measure. Second, different international inequality databases frequently produce different results, even when the countries, the welfare concept, the inequality measure, and the time period are held constant. Third, while there is a rather large amount of evidence that estimated rates of convergence differ by region and by time, even this result is sensitive to the database that is used to perform the analysis.

Essays on Public Policy and Economic Inequality

A DISSERTATION

SUBMITTED ON THE EIGHTEENTH DAY OF APRIL, 2017

TO THE DEPARTMENT OF ECONOMICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

OF THE SCHOOL OF LIBERAL ARTS

OF TULANE UNIVERSITY

FOR THE DEGREE

OF

DOCTOR OF PHILOSOPHY

BY

Daniel Teles

APPROVED:

Alan Barreca, Ph.D.
CHAIR

Stefano Barbieri, Ph.D.

Steven Sheffrin, Ph.D.

Acknowledgements

I was fortunate to have the benefit of insightful comments from many friends and colleagues as this dissertation progressed. I am especially grateful for the guidance from my dissertation committee Alan Barreca (chair), Stefano Barbieri, and Steven Sheffrin who provided valuable feedback at every stage and to Nora Lustig who co-authored chapter 3 while providing mentorship and support.

Thanks are due to Bibek Adhikari, Jon Bakija, Whitney Ruble Bross, Grant Driessen, Paige Miller, Maggie Mullane and a number of anonymous reviewers, all of whom read various drafts of these essays, and to Patrick Button and Wei Long, who provided excellent advice and econometric insight. I also received helpful feedback at conferences held by the Corporation for National and Community Service, the National Tax Association, the Society of Government Economists, and the Southern Economic Association.

This research would not have been possible without funding from the Corporation from National and Community Service and the Murphy Institute, nor without the support of Agnieszka Nance, Bridget Smith, and the staff at the Center for Public Service. Both credit and thanks are owed to my research assistants, Allison Koh, Mengde Liu, Sarah Margolies, and Lu Yao who handled a great deal of the data cleaning necessary to make the analysis in the second chapter of this dissertation possible.

Most importantly, I want to thank Lauren Gulkus, for everything else.

Contents

List of Tables

List of Figures

Introduction

This dissertation consists of three essays, each of which has implications for public policy. My research interests began with curiosity about the impacts of public policy on the economically disadvantaged and diverged along two strands of economic research. The first two essays estimate the economic effect of policies intended to support the nonprofit sector. These essays are greatly informed by the literature on the private provision of public goods. The third essay examines the sensitivity of research on income inequality to various sources of data. The impetus for this essay, which was be written jointly with Professor Nora Lustig, is the growing use of inequality metrics in both domestic and international policy evaluations.

I begin with an analysis of tax credits that encourage charitable giving. The majority of U.S. states have implemented some form of tax incentives to encourage individuals or businesses to donate to specific classes of nonprofits. However, the impact of these policies on the nonprofits that they are designed to support has been unknown. My analysis focuses on two of the costliest programs in terms of lost tax revenue. The WPTC, the largest tax credit for charitable giving in terms of tax expenditure, provides a broadly targeted 100 percent credit with a cap of $200 per person during the sample period. Endow Iowa provides a sharply targeted 25 percent credit with a cap of $300,000 per taxpayer. I include in the essay a model of donor budget constraints and preferences and hypothesize that the construction of Endow Iowa should induce substitution effects that benefit community foundations in Iowa. Using synthetic control methods to construct counterfactuals, I estimate a 125 percent increase in contributions to community foundations in Iowa. Moreover, the growth in contribution levels in Iowa included increases in both the number of community foundations and the level of contributions per foundation. In contrast, I find no evidence that the WPTC increased contributions to the targeted Arizona nonprofits.

My second essay examines the economic impact of the AmeriCorps program on the nonprofits. AmeriCorps is a network of service programs throughout the United States that embeds its members, among other places, in 501(c)3 nonprofits. I examine the AmeriCorps program through the lens of a growing economic literature on the interaction between private charitable giving and government funding for public goods. In its analysis of the relationship between the AmeriCorps program and private donations to the nonprofits with which it partners, this is the first analysis of whether government funded labor might "crowd out" or "crowd in" private capital. I separately estimate crowd out and crowd in effects for the AmeriCorps State and National and AmeriCorps VISTA programs, and along the extensive and intensive margins. Using a fixed effects model, I find that both AmeriCorps types are associated with higher levels of private contributions. However, an instrumental variables (IV) analysis finds a negative causal effect, suggestive of crowding out. The IV analysis shows a small, statistically insignificant impact of VISTA on contributions but finds that a ten percent increase in the number of AmeriCorps State and National is associated with a one percent decline in contributions. This level of crowd-out is similar to those estimated for other forms of government funding for the nonprofit sector. The unique structure of AmeriCorps does not appear to have a unique effect nonprofit revenue.

The final essay tests the sensitivity of cross-national inequality research to the choices about the underlying data. To do this, the essay examines the extent to which estimates of inequality convergence are sensitive to the choice of welfare concept, inequality indicator, database, country coverage, and time period. Moreover, we explore the sensitivity of the estimated rate of convergence by testing five hypotheses using a series of pair-wise F-tests. The main takeaways are as follows. First, estimates appear to be more sensitive to the choice of welfare concept than to the choice of inequality measure. Second, different international inequality databases frequently produce different results, even when the countries, the welfare concept, the inequality measure, and the time period are held constant. Third, while there is a rather large amount of evidence that estimated rates of convergence differ by region and by time, even this result is sensitive to the database that is used to perform the analysis.

Chapter 1

Do Tax Credits Increase Charitable Giving? Evidence from Iowa and Arizona

1.1 Introduction

The nonprofit sector represents a private alternative to government provision of public goods. Institutions such as universities, hospitals, and museums are provided by both governments and nonprofits, while nonprofit charities supplement the work of government to decrease homelessness, cure diseases, and broaden economic opportunity. Whereas increases in government production must be funded with tax revenues, nonprofit production draws voluntary contributions from donors. It is therefore not surprising that policymakers often look to the nonprofit sector to increase the well-being of their constituents without increasing their tax burden. In the United States, a popular policy option at the state level has been the introduction of tax credits that encourage philanthropic giving that is directed toward a specific area of need.

Charitable tax credit (CTC) programs provide incentives to taxpayers who donate to either pre-approved or specific classes of nonprofits. De Vita and Twombly (2004) describe the policy goals of CTC programs as follows: "(1) to increase charitable giving, (2) to allow taxpayers to determine directly the utility or effectiveness of charitable services, and (3) to support antipoverty programs," (p. 1). Generally, CTC programs also require that donations are made to a nonprofit within the state in order to qualify for the credit. This additional goal is described, in a report on the Endow Iowa Credit, as reducing "the transfer of wealth to outside of the state," (Gullickson and Tilkes, 2013, p. 10).[1]

Research into the effectiveness of tax credits, of various types, continues to generate interest. For example, a notable strand of literature finds that firms respond to tax credits for research and development (Bloom et al., 2002; Lokshin and Mohnen, 2012). Additional research into the efficacy of tax credits as policy instruments finds that firms increase employment when provided with credits designed to create jobs (Faulk, 2002) or encourage innovation clusters (Moretti and Wilson, 2014). The response of individuals appears to be more nuanced. For example, Ramnath (2013), finds that credits designed to increase retirement savings led individuals to manipulate their income to take advantage of the credit but does not find statistically significant evidence of increased savings. The effectiveness of tax credits designed to spur donations and grow the nonprofit sector, however, has not been the subject of rigorous research.

The responsiveness of donors to changes in the after-tax price of giving has been well researched, but the effect on recipients (the charities themselves) has been largely ignored.[2] In the existing research, the relationship between the changes in the after-

[1] Gullickson and Tilkes (2013) provides a detailed description of the Endow Iowa program, presents findings on donor demographics, and estimates that making an Endow Iowa-qualifying donation is associated with an additional $0.09 more in non-Endow Iowa contributions.. Causal estimates were beyond the scope of their study.

[2] Recent work by Duquette (2016) provides an exception, examining the effect of the federal tax deduction for charitable giving on nonprofit recipients. This paper differs from Duquette (2016) by focusing on state level tax credits.

tax price and the quantity of charitable giving is generally summarized by a price elasticity. A number of studies have estimated the elasticity relative to the U.S. federal deduction for charitable giving (Auten et al., 2002; Barrett et al., 1997; Randolph, 1995). Since the tax deduction treats all giving equally, all donations have an after-tax price of one minus the individual's marginal tax rate. Variation in marginal tax rates across time and across individuals can then be used to identify tax-price elasticity. By including variation in state income tax policy in their analysis, Bakija and Heim (2011) separately identify federal and state tax-price elasticities. In all of these studies, the price of charity varies by individual but does not vary by recipient. As such, these studies do not predict the effect of tax-credit polices that intentionally make giving to certain nonprofit organizations cheaper than giving to others.

Anecdotal evidence suggests that targeted charitable tax credits have large effects on the revenue streams of nonprofit organizations. In St. Louis, The Hunger Task Force of the Missouri Association for Social Welfare reported that a food pantry in St. Louis lost more than $30,000 in donations after a credit for giving to food pantries expired in 2011 (National Council of Nonprofits, 2015). In Detroit, the Coalition on Temporary Shelter (COTS) reported a 10 percent decline in charitable contributions following the repeal of Michigan's Homeless Shelter/Food Bank tax credit. A Grand Valley State study of charitable giving in Michigan found that donations of less than $400 fell between 2011 and 2012, the year in which the tax credits were removed (Johnson Center for Philanthropy, 2013). The results of these studies, however, did not take into account other economic factors at play and therefore were unable to provide convincing evidence of a causal relationship.

This study focuses on two of the largest CTC programs in terms of tax expenditure: the Endow Iowa Tax Credit ("Endow Iowa") and Arizona's Working Poor Tax Credit (WPTC). The programs provide stark contrast in their structures, and both are well suited to rigorous analysis. While Endow Iowa provides a sharply targeted

25 percent credit with a cap of $300,000 per person, the WPTC provides a broadly targeted 100 percent credit that was capped at $200 per person during the sample period. The cost of each program is well documented; in each case, the state's Department of Revenue has tracked annual tax expenditure. The level of tax expenditure associated with each credits is large relative to similar programs in other states. Most importantly, both have clear regulations and public lists of qualifying organizations, which make it possible to identify appropriate treatment and control groups.

A model of the budget constraints and preferences of donors suggests that Endow Iowa has the potential to induce large increases in contributions to community foundations. Donors with Cobb-Douglas preferences increase contributions to community foundations by one third. In the extreme case, some donors would shift most or all of their giving away from other charities and toward community foundations. The framework of the WPTC, conversely, has only an income effect. As such, the model predicts that donors respond by increasing donations to targeted nonprofits by no more than $200.

This paper further differs from the existing literature by performing a rigorous, empirical analysis of the effect of targeted (rather than across-the-board) tax policies on the recipients of charitable donations (rather than on the donors). Specifically, I examine whether Endow Iowa and the WPTC increased the amount of contributions received by the targeted nonprofits. The greatest hurdle in answering this question is the development of appropriate counterfactuals. I estimate counterfactuals using the Synthetic Control Methods (SCM) described in Abadie and Gardeazabal (2003), Abadie, Diamond, and Hainmueller (2010), and Abadie, Diamond, and Hainmueller (2015). The methodology compares aggregate data for a treatment group to a synthetic, untreated version of itself.

I find that contributions to community foundations in Iowa far exceeded those of a synthetic control, suggesting that the Endow Iowa tax credit, at least in conjunction with the related County Endowment Fund Program (CEFP), led to higher contribution levels. There is evidence of a large increase in donations—more than \$45 million in additional annual contributions stemming from \$2 to \$6 million in credits and \$6 to \$11 million in grants. The primary result holds under a variety of alternative specifications for the synthetic control. Additional analysis is suggestive of both a growth in the number of community foundations and in increases in per-foundation contribution levels.

I find no evidence that Arizona's WPTC increased charitable contributions levels. I estimate the gross change in contributions by focusing on a non-random group of nonprofits that received high levels of qualifying donations. In my baseline specification, the aggregated group received lower levels of donations than their synthetic control.

The remainder of this paper is structured as follows. Section 2 provides a brief overview of state tax credit policy, a theoretical framework for understanding the interaction between tax policy and contribution levels, and proposes testable hypotheses. Section 3 then describes the methodological framework used in the primary analysis, including discussion of the underlying data, Synthetic Control Methods, and identification of the treatment and control groups. Section 4 presents the primary results for Iowa and Arizona. I include extensions in Section 5 and a discussion of the findings in Section 6.

1.2 Tax Credits and the Price of a Donation

1.2.1 Charitable Tax Credit Policies

In the United States, a majority of state legislatures have provided some form of CTC. In 2013, 33 states offered at least one CTC program to businesses, and 20 states offered charitable giving tax credits to individuals. While there are some similarities between programs there are a number of dimensions upon which they vary. Endow Iowa and the WPTC in Arizona are among the largest programs in terms of tax expenditures. Endow Iowa provides an example of a highly targeted 25 percent credit with a large cap on eligible donations. In contrast, Arizona's WPTC, provides an example of a broad-based, 100 percent credit with a low cap. I model these differences in Section 1.2.2.

Table 1.1 summarizes differences between 14 CTC policies.[3] Three important variations are the size of the credit as a percentage of the donation, the statutory credit cap (the maximum amount any one taxpayer may receive), and the nonprofits targeted by the credit. Credits range from 15 percent to 100 percent of the qualifying donation. The cap on the credits, meanwhile, ranges from a low of $100 for individuals (or $200 for couples) to a high of $300,000. The largest program in terms of expenditure is the WPTC, which cost the state $21.8 million in 2012. The smallest programs are the Endow Kentucky Program and Nebraska's Qualified Endowment Credit. Part of this difference in total expenditure lies in the type of giving the programs plan to stimulate. Programs such as Arizona's Working Poor Tax Credit (WPTC) and Michigan's Homeless Shelter and Food Bank Credit appear to be geared toward small donors making regular contributions to local nonprofits. Conversely, Endow Iowa, the Neighborhood Assistance Tax Credits in Connecticut and Delaware, and Missouri's Youth Opportunities Program focus on larger, one-time gifts.

Along with being among the largest CTC programs, Endow Iowa and the WPTC are well documented. As such, they are well-suited to rigorous analysis. The annual tax expenditures associated with both Endow Iowa and the WPTC were closely tracked and reported by the states' Departments of Revenue, along with detailed summaries of policy regulations and regulatory changes (Gene, 2013; Gullickson and Tilkes, 2013). The provisions of both the Endow Iowa Credit and the WPTC stip-

[3] The 14 credit programs included in Table 1.1 are those for which I was able to obtain tax expenditure information.

ulate that only donations made to nonprofits that have pre-qualified with the state are eligible. The documentation of these pre-qualified nonprofits allow for the identification of treated organizations and comparison groups in other states. While the WPTC program is generally broad, I focus on a group of qualifying organizations for whom credits were equal to two percent of their charitable receipts. Endow Iowa represents an even greater share of contributions. Since its inception, the program has distributed credits equal to three percent of total contributions to community foundations.

Endow Iowa provides a sharply targeted credit of up to $300,000 per taxpayer. The credit program went into effect January 1, 2003. At the time, the credit was equal to 20 percent of donations made to any permanently endowed fund within a *qualified foundation*. Rather than directly providing public goods and services, foundations pool donations into a coordinated investment fund from which grants are given to other nonprofit charities. Foundations must apply to the Iowa Economic Development Authority to qualify for the program. Initially, taxpayers were allowed to take both the new credit and the preexisting deduction. Beginning in 2010, the credit was increased to 25 percent, but taxpayers were no longer allowed to include their donation as part of itemized deductions. The total amount of credits received by all taxpayers is capped each year, beginning at $2 million for all of 2003 and 2004, but rising to $6 million as of 2012 (Gullickson and Tilkes, 2013).

A year after its implementation, Endow Iowa was followed by the introduction of the CEFP. The CEFP distributes a small percentage (0.5 percent in 2004 and 2005, an 0.8 percent thereafter) of state gaming tax revenue to community foundations associated with counties that do not have gaming licenses (Iowa Council of Foundations and Iowa Gaming Association, 2013). The program distributes the money annually, with the first distributions—funded with 2004 tax collections—disbursed in 2005. In total, $92.7 million has been distributed to community foundations (Iowa Department of Revenue, 2014).

The WPTC provides a broadly targeted tax incentive that was capped at $200 per individual until 2013. First implemented in 1998, the policy provides a non-refundable income tax credit for cash contributions to qualifying charitable organizations. Qualifying organizations are defined as those that spend at least 50 percent of their budget on Arizona residents who receive Temporary Assistance for Needy Families (TANF) benefits, have a household income less than 150 percent of the poverty level, or have chronically ill or disabled children. This definition allows contributions to a wide range of charities—809 nonprofits qualified in the 2009 tax year (Gene, 2013). Initially, the credit only applied to donations in excess of a baseline amount donated in a base-year (generally the year before the individual first applies for the credit) and to individuals who itemized their deductions. The credit was equal to 100 percent of the first $200 in excess contributions or $400 for married couples filing jointly (Gene, 2013). Arizona altered the WPTC rules in 2009 and again in 2013. Beginning in 2009, the establishment of a baseline donation was removed, and individuals were allowed to apply the credit to their first $200 of contributions. Legislation in 2013 expanded the credit to non-itemizers and increased the cap for donations to foster care organizations to $400, or $800 for married couples filing jointly (Gene, 2013).

The WPTC and Endow Iowa were designed very differently and therefore make interesting points of comparison. The WPTC distributes a large number of credits (49,915 in 2009) in relatively small amounts—an average of $272 per claim in 2009 (Gene, 2013). Conversely, Endow Iowa distributes fewer credits (3,074 in 2012) but awards an average credit of $1,884, with more than 70 percent total tax-savings accruing to individuals who gave more than $30,000. (Gullickson and Tilkes, 2013). Arizona's program qualifies a broad array of nonprofits, while Iowa focuses directly on community foundations. The specificity of Endow Iowa may make estimating its impact easier. However, because of the breadth of the WPTC, it includes national organizations such as Big Brothers Big Sisters and Habitat for Humanity, which can be compared to their counterparts in other states. Endow Iowa is available to businesses as well as individuals, while the WPTC is not. Most interestingly, the programs differ in the percentage of donation credited (25 percent in Iowa and 100 percent in Arizona) and in the credit cap ($300,000 in Iowa and $200 in Arizona). These two differences can be explored further by investigating their effects on donors' budget constraints.

1.2.2 A Framework for Analysis

A primary goal of any charitable tax credit program is "to increase charitable giving" (De Vita and Twombly, 2004, p. 1). This is accomplished by lowering the after-tax price of a donation. The federal tax deduction for charitable gifts operates on the same premise, but it does not target a specific subset of the nonprofit sector. By making some donations cheaper than others, targeted CTCs produce a substitution effect *between charities*, which is modeled below. The model displays effects of Endow Iowa and the WPTC on both targeted and untargeted charities. CTCs may also alter the return to fundraising expenditures by nonprofits; however, the impact of this alteration is ambiguous. In Iowa, the CEFP, too, may influence contribution levels. Fortunately, existing research provides estimates that can be used to bound the its impact.

The model examines donations to two nonprofit charities, indexed 1 and 2, and assumes that there exists a predetermined total amount that some individual wishes to contribute to charity, $\bar{G}$. Setting a predetermined level of charitable giving for each consumer is a simplifying assumption imposed upon the model in order to focus on substitution between charities rather than between charitable giving and private consumption. Implications of relaxing this assumption are discussed following derivation of the model. Implicitly, this assumption defines a price elasticity for total charitable giving of -1 and defines donor motivations as consisting entirely of warm-glow .[4]

A price elasticity of -1 for total charitable giving implies that donors adjust their contributions proportionally to changes in the price of donations. As such, the increase in total charitable gifts is exactly equal to the value of the tax incentives awarded by the government. Moreover, the assumption that charitable donations are unit elastic has several other important implications: (1) cross-price elasticity on private consumption is 0; (2) the income effect of tax incentives for charitable giving on private consumption is equally offset by substitution into charitable giving; and (3) private consumption is fixed. This assumption corresponds to existing empirical estimates of the price elasticity of charitable giving, which generally cluster around -1 and include both the income effect of a change in the after-tax price of donations and substitution into charitable giving.[5]

The assumption that donors are entirely motivated by warm-glow excludes equilibrium effects from the model. I assume that donors care only about the value of their own gift and that neither the total level of contributions nor the output of the charity enters their utility functions. Therefore, no substitution between private consumption and charitable giving occurs in response to changes in the total level of revenue at either charity 1 or 2.[6]

To begin, let an individual's donation set be defined $G = (g_1, g_2)$ and define her

[4] Warm-glow refers to the utility gains from the value of the gift and is often contrasted with altruism, or the utility gains from a public good produced by a government or a charity. Andreoni (1990) formalizes a model in which donors may be motivated by either warm-glow, altruism, or both.

[5] Auten, Sieg, and Clotfelter (2002) estimate elasticities between -0.79 and -1.26. A meta-analysis by Peloza and Steel (2005) reports that the weighted average of prior studies is -1.44 or -1.11 with outliers removed. More recently, Bakija and Heim (2011) separately estimate tax-price elasticity for changes in federal and state taxes, finding an elasticity of -1.16 at the state level.

[6] The model also ignores the potential costs of learning about and filing for CTCs. In Iowa, community foundations generally help their donors register for the credits, for which they receive a certificate number from the Iowa Economic Development Authority. In Arizona, donors need only to complete form 321 and attach a receipt. Again, nonprofits are generally forthcoming with assistance, many provide information on-line and in fundraising documents, and some—e.g. Habitat for Humanity of Central Arizona and St. Vincent De Paul of Southern Arizona—even include a PDF of form 321 on their websites. I therefore assume that, at least among individuals who are considering donations to targeted nonprofits, the additional costs of applying for a credit are minimal.

budget constraint $\bar{G} = p_1 g_1 + p_2 g_2$, where p_1 and p_2 are the after-tax prices of giving to charity 1 and 2, respectively. In the absence of incentives, the price of giving is equal to one; the budget constraint is defined as $\bar{G} = g_1 + g_2$ and represented as a line with the slope -1. A diagram of this budget constraint appears in the top left panel of Figure 1.1. Tax incentives alter the prices of donations g_1 and g_2 by offering a subsidy to donors. For a given subsidy rate s_j, the after-tax price is $p_j = 1 - s_j$. An uncapped, untargeted tax incentive sets some price $p_u = p_1 = p_2 \leq 1$. The new budget constraint, which appears in the top right panel of Figure 1.1, is defined as $\bar{G} = p_u(g_1 + g_2)$. The federal deduction for charitable giving is an untargeted incentive of this type with s_j equal to one minus the taxpayer's marginal tax rate. Targeted tax credit programs only change the price of the donation to the targeted nonprofit, g_1, leaving $p_2 = 1$. An uncapped, targeted program would therefore lead to the following budget constraint

$$\bar{G} = p_1 g_1 + g_2 \tag{1.1}$$

with $p_1 \leq 1$ and a slope of $-p_1$.

A credit program with a cap creates a kink in the budget constraint, as shown in the bottom two panels of Figure 1.1. Defining the credit cap as C, the kink occurs when the donor has contributed $g_1 \geq C/(1 - p_1)$.[7] Therefore, the budget constraint for a targeted program with a cap can be written as

$$\bar{G} = \begin{cases} p_1 g_1 + g_2 & \text{if } g_1 \leq \frac{C}{1-p_1} \\ -C + g_1 + g_2 & \text{if } g_1 \geq \frac{C}{1-p_1} \end{cases} \tag{1.2}$$

Up to the cap, the slope of the budget constraint is, $-p_1$. Beyond the cap, the slope is -1 and the budget line is parallel to the initial budget constraint.

The budget constraints of donors in Arizona and Iowa are modeled in the bottom

[7] The credit cap is the maximum amount the government is willing to distribute to an individual as a tax incentive. For some subsidy rate s_1, the donor can only receive subsidy until $s_1 g_1 = C$. Given $p_1 = 1 - s_1$, it follows that the cap is reached when $C = (1 - p_1)g_1$ or $g_1 = C/(1 - p_1)$.

panels of Figure 1.1. An individual in Arizona may donate up to \$200 to the targeted nonprofits and receive a 100 percent credit. With a 100 percent credit, $p_1 = 0$. A credit of this type shifts the budget constraint to the right but does not alter the relative price of giving beyond the cap. For the first \$200, the donor pays a price of zero and the budget constraint is represented as a horizontal line. Above the cap, the tax credit is no longer applied, so $p_1 = p_2$ and the budget constraint is represented by a line with slope -1. In contrast, Endow Iowa alters the relative price of giving to the targeted nonprofits. A 25 percent credit leads to an after-tax price of 0.75. Endow Iowa's cap is \$300,000, and therefore the kink in the budget constraint does not occur until a donor has contributed \$1.2 million to a community foundation. Figure 1.1 shows the effect of Endow Iowa on the budget constraint of a donor with $\bar{G}$ equal to \$1.2 million dollars.

To complete the model, assume that for all donors charity 1 and charity 2 are normal goods. Let G_0 represent the donation set in the absence of a tax incentive, and G^* represent the donation set in the presence of a CTC program. Additionally, assume that for each donor the total donation budget $\bar{G}$ is between \$201 and \$1.2 million so that they may spend beyond $C/(1-p_1)$ in Arizona but not in Iowa.[8] For any given donor's preferences, it is now possible to examine the impact that Endow Iowa and WPTC would have on donations to the targeted and untargeted charitable sectors. I summarize the expected response of donors with different preferences in Table 1.2 and illustrate three preference types in Figure 1.2.

Some donors may see the charities as perfect substitutes but, when prices are equal, favor one type over another. Their utility can be defined as $U = \phi_1 g_1 + \phi_2 g_2$. If $\phi_1 > \phi_2$, the donor prefers targeted charities in the absence of a credit, and if $\phi_1 < \phi_2$, the donor prefers the untargeted charities in the absence of a credit. These

[8] A contribution of \$1.2 million far exceeds the size of a typical donation in Iowa. From 2005 to 2012, only 13 donors qualified for the maximum credit award (Gullickson and Tilkes, 2013). As such, Iowa's cap can be treated as not binding and can be excluded from the model.

preferences are depicted in the top and bottom rows of Figure 1.2.

With Endow Iowa in place, donors who would have preferred the targeted charities (community foundations) in the absence of the credit would increase contributions to the targeted sector by one third, as depicted in the top right panel of Figure 1.2. Donors who would have preferred the untargeted charities behave differently based on the relative strength of their preference. If $\phi_1 < \phi_2$ and $\frac{\phi_1}{\phi_2} < 0.75$, the donor prefers the untargeted charities and does not respond to a 75 percent credit. The policy would have no effect on her. If $\phi_1 < \phi_2$ and $\frac{\phi_1}{\phi_2} > 0.75$, the donor prefers the untargeted charities and shifts *all of her contributions* to the targeted charities. This scenario appears in the bottom right panel of Figure 1.2. In contrast, the WPTC has the same effect on donations for any positive levels of ϕ_1 and ϕ_2. In either case, these donors are expected to respond to the WPTC by increasing their donations by exactly the \$200 of the cap. As such, $G^* - G_0 = \$200$ whether G_0 is positive or zero.

The middle row of Figure 1.2 shows how a donor with Cobb-Douglas preferences of type $U = g_1^{\alpha} g_2^{\alpha-1}$ would respond to an untargeted incentive, the WPTC, and Endow Iowa. In the absence of a credit, this type of donor would donate to both the targeted and untargeted nonprofits. Since the price of giving in Arizona is zero when giving less than the cap, these donors can be expected to donate at least \$200 to the targeted nonprofits. Any donor who would have contributed that amount even in the absence of the credit sees the WPTC as a lump sum transfer. In this model, she would then increase total giving by the value of the credit with contributions increasing such that the proportion of contributions going to targeted nonprofits is unchanged. Donors with Cobb-Douglas preferences in Iowa would behave in this manner only if they are contributing in excess of \$1.2 million to community foundations. Otherwise, their budget and preferences are described by the right-middle panel of Figure 1.2. These donors increase their contributions to targeted nonprofits by one third and leave donations to untargeted nonprofits unchanged.

Finally, suppose donors exhibit constant elasticity of substitution (CES) type utility with preferences defined $U = (\phi g_1^{\frac{\epsilon-1}{\epsilon}} + (1-\phi)g_2^{\frac{\epsilon-1}{\epsilon}})^{\frac{\epsilon}{\epsilon-1}}$. Donors then see contributions to targeted and untargeted charities as gross complements when $\epsilon < 1$ and as gross substitutes when $\epsilon > 1$. Additionally, they prefer the targeted charities when $\phi > 0.5$ and prefer the untargeted charities when $\phi < 0.5$. When the charities are perceived as complements, the donors respond to tax credits by increasing spending on contributions to both targeted and untargeted nonprofits. As ϵ approaches zero, the two charity types become perfect complements. Donors in Arizona who see targeted and untargeted nonprofits as perfect complements increase giving to each by \$100. Donors in Iowa who see community foundations and other nonprofits as perfect complements increase giving to each by 14 percent. As ϵ grows, CES preferences approach the linear substitutes displayed in the top and bottom rows of Figure 1.2. In Iowa, large substitution effects occur when donors both are relatively indifferent between the targeted an untargeted charities ($\phi \approx 0.5$) and see the charities are close substitutes ($\epsilon \gg 1$). Since the WPTC acts as a lump sum transfer, no substitution effects occur even in the extreme case.

What happens if the assumption that consumption is unaffected by incentives for charitable giving is relaxed? Begin by defining the budget constraint as a function of income M rather than $\bar{G}$. In the absence of a tax incentive, the budget constraint is defined as $M = g_1 + g_2 + x$, where x is the total expenditure on consumption. Targeted tax credit programs with a cap on credits produce the following budget constraint,

$$M = \begin{cases} p_1 g_1 + g_2 + x & \text{if } g_1 \leq \frac{C}{1-p_1} \\ -C + g_1 + g_2 + x & \text{if } g_1 \geq \frac{C}{1-p_1} \end{cases} \tag{1.3}$$

Below the cap, we expect contributions to the targeted nonprofit to increase, but the amount is ambiguous. If the substitution effect between consumption and charity dominates, consumers decrease both donations to charity 2 and their private consumption and increase contributions to the targeted nonprofit.

This additional substitution from private consumption would lead to larger increases in contributions to the targeted nonprofit than in the baseline model. Moreover, if we assume that the marginal warm-glow utility from donating declines as the donor increases her contribution, substitution from private consumption into charitable giving could mitigate the decline in contributions to the untargeted charities that the model predicts. The more a donor substitutes away from her consumption spending and toward gifts to charity 1, the lower the marginal utility from additional gifts to charity 1 and thus the smaller the substitution effect from charity 2 to charity 1.

If income effects dominate, consumers increase donations to both charities and their consumption expenditures such that the increase in contributions to the targeted nonprofit is less than the value of the credits and therefore less than in the baseline model. Note that it would also be possible for contributions to charity 2 to rise while consumption expenditure fell, or vice versa.

As in the base model, consumers who would contribute an amount greater than $\frac{C}{1-p_1}$ in absence of the subsidy see no price change and the substitution effect is zero. As such, there should not be a decline in either donations to untargeted charities or consumption expenditures. Consumers receive the credit as a lump sum transfer and distribute it to maximize their utility. If donors increase private consumption, the predicted increase in contributions to both the targeted and untargeted nonprofits is smaller than in the baseline model (without adjustment).

This framework does not include endogenous changes in fundraising. If individual donations increase as expected, nonprofits get a better return for each dollar spent soliciting a new donor. Revenue-maximizing charities, such as those modeled by Name-Correa and Yildirim (2013), would therefore increase their fundraising expenditures.[9] However, there is evidence that at least some nonprofits spend below the revenue maximizing level on fundraising (Khanna et al., 1995; Okten and Weisbrod, 2000). Without evidence to suggest that the nonprofits targeted by CTCs are revenue maximizers, the direction of endogenous fundraising change is ambiguous.

Analysis of Endow Iowa is further complicated by the presence of its sister policy, the CEFP. Fortunately, the literature on the interplay between government grants and private giving can illustrate the extent to which the grant program is likely to confound the estimated impact of Endow Iowa. Models of altruistic giving (donors motivated by public good production), such as Bergstrom, Blume, and Varian (1986), suggest that government spending on public goods should crowd out (reduce) private giving. Conversely, Payne (2001) explains that if donors possess imperfect information, government grants can serve as a signal of quality and thus crowd in (increase) donations. The model presented here, which assumes donors are motivated exclusively by warm-glow, suggests that donations are unaffected by government grants to charity.

Empirical estimates of crowd out help bound the likely impact of the CEFP. Tinkelman (2010) summarizes 46 empirical studies on the impact of grants on private giving and finds that, while estimates vary, they center around zero and only eight include specifications in which crowd out is estimated to be greater than 50 percent. If the effect of the CEFP is similar to the grant programs that have been studied in the past, the result of the program would be an increase in donations of between 50 and 150 percent of the value of the credits—$4.9 to $14.6 million dollars annually between 2004 and 2012. If the effect falls toward the middle this distribution, the CEFP would increase revenue levels by exactly the value of the grants and, as in the model described above, private contributions would be unaffected by the program.

[9] Note, however, that the Name-Correa Yildirim model was not derived to examine the impact of a change in the tax-price of donations.

1.2.3 Hypotheses

The primary hypothesis of this study is as follows.

Hypothesis 1 *Charitable Tax Credit programs increase the level of contributions received by targeted nonprofit organizations.*

This hypothesis is born from stated policy goals and the extensive empirical literature surrounding tax deductions for charitable giving. De Vita and Twombly (2004) note that CTC programs share the common goal of increasing charitable giving, and Gullickson and Tilkes (2013) state that Endow Iowa began "in an effort to spur philanthropy across Iowa" (p. 10). Moreover, existing studies of donor responses to tax incentives—e.g., Auten et al. (2002), Peloza and Steel (2005), and Bakija and Heim (2011)—have found strong evidence that tax incentives increase charitable giving.

The framework presented in 1.2.2 leads to four additional hypotheses (two for each policy).[10] First as illustrated in the middle frame of Figure 1.2, the WPTC is likely to lead to positive spillover, increasing donations to untargeted nonprofits. Second, as illustrated in the bottom right frame of Figure 1.2, Endow Iowa may lead to substitution between charities and negative spillover, reducing donations to untargeted nonprofits. Third, the increase in contributions to targeted nonprofits in Arizona should be no greater than the value of the credits. Fourth, the total increase in contributions to community foundations in Iowa will be greater than the cost of the tax credits awarded. This occurs because the increase in contributions induced by Endow Iowa will be equal to the sum of the credits received by donors and the substitution from charities that do not qualify for credits to those that do.

While these hypotheses suggest contrasts between Iowa and Arizona, I do not formalize any hypotheses about direct comparisons. The model described in the

[10] As in the baseline model, these hypotheses assume that donors have a fixed budget for charitable giving, a price elasticity for charitable gifts of -1, are motivated entirely by warm-glow and see charitable gifts as normal goods. Additionally it is assumed that there are no equilibrium effects induced by changes in fundraising behavior by charities.

previous section allows for the analysis of the alternative subsidy rates and credit caps proposed in Iowa and Arizona. However, it ignores the important difference in the type of nonprofit organization targeted by the credit. The model would allow for direct comparison between credit programs if the two charity types and donor preferences were held constant across states. Since the two programs target different sectors, however, this is not the case. Even where the existing hypotheses suggest comparison relative to the value of the credits awarded, as is the case with Hypothesis 4 and 5, I do not formalize a hypothesis because the synthetic control methodology, described in Section 1.3, does no provide an appropriate empirical test.

Under the model, donors who contribute enough to receive the full credit cap see no change in prices and respond to the subsidy as if it were a lump sum increase in income, leading to the following hypothesis.

Hypothesis 2 *Arizona's WPTC program leads to positive spillover, increasing donations to untargeted nonprofits.*

The model shows that donors who see charities as perfect substitutes will increase their donations to the targeted nonprofit, g_1, by the amount of the credit and leave their donations to charity 2 unchanged. This holds true whether or not they would have donated to the targeted nonprofit in the absence of a credit (that is, whether they have preferences represented in the top or bottom row of Figure 1.2). For all other donors, we should expect to see an increase in donations to both targeted and untargeted nonprofits. The expected result is therefore an increase in total contributions by the amount awarded in credits, with some level of positive spillover for untargeted nonprofits from donors who do not see the charities as substitutable. Since the model assumes a fixed budget for charitable giving, the total increase in donations must be equal to the value of the tax credits. In Arizona, the model predicts that some of these additional resources are spent on untargeted charities. Therefore, the increase in contributions to targeted nonprofits in Arizona should be less than the value of the credits. If this assumption were relaxed, the spillover effects could extend to non-charitable expenditures. However, we would still not expect to find evidence of substitution from untargeted to targeted nonprofits.

When CTC subsidies are greater than zero and less than 100 percent and the donors' budget does not allow them to reach the credit cap, there is the potential for large substitution effects. The model shows that, below the cap, donors who see contributions to community foundations and other nonprofits as complementary will increase contributions to both, and donors who see donations as gross substitutes will tend to shift their spending away from untargeted nonprofits and toward community foundations. Given that the average donation to community foundations in Iowa is far below the credit cap (Gullickson and Tilkes, 2013), the model predicts that if the average donor sees donations to community foundations and other nonprofits as gross substitutes, the policy would induce a decline in contributions to untargeted charities. This finding leads to the following hypothesis.

Hypothesis 3 *The Endow Iowa tax credit produces a large substitution effect between charities and reduces donations to untargeted charities.*

Support for Hypothesis 3 could occur even if a majority of donors do not see the two goods as gross substitutes. The model predicts that the increase in contributions by donors who see the charities as complements can be no greater than the value of the subsidy received by the donor. In contrast, donors who see the two charities as gross substitutes may decrease their donations to the untargeted charity by up to 100 percent. In the extreme case in which charities are perceived as perfect substitutes, donors may switch from giving exclusively to untargeted nonprofits to giving exclusively to targeted nonprofits. A large effect could be driven by a small number of donors who only began giving to community foundations after the introduction of Endow Iowa and decreased contributions to untargeted nonprofits (almost) entirely. Conversely, rejection of Hypothesis 3, and the null that there is no change in contributions to untargeted nonprofits, supports the conclusion that the average donor sees contributions to the two charity types as gross complements.

Under the assumption that private consumption is unaffected by a change in the price of donations, the total increase in charitable giving to targeted nonprofits is equal to the value of the tax credits awarded plus (minus) the reduction (increase) in contributions to untargeted charities.[11] As such, Hypotheses 2 and 3, respectively, imply the following suppositions:

Hypothesis 4 *The WPTC leads to an increase in charitable giving to targeted non-profits that is less than the value of credits disbursed.*

Hypothesis 5 *Endow Iowa leads to an increase in charitable giving to community foundations that is greater than the value of credits disbursed.*

If the price elasticity of charitable giving is -1, and existing research suggests that this is (at least approximately) the case, any additional money given to the taxpayer as a subsidy for donations will be passed on in the form of donations. In Arizona, we expect these funds to be divided between targeted and untargeted charities. If, however, Endow Iowa induces substitution of donations away from untargeted charities and toward targeted charities, the increase in donations to the targeted sector must be larger than the value of the credits.

[11] If that assumption is relaxed, the total increase in charitable giving is equal to to the value of the tax credits awarded plus (minus) the value of contributions shifted between untargeted and targeted charities plus (minus) any decrease (increase) in private consumption.

1.3 A Case-Study Approach Using Synthetic Control Methods

1.3.1 Data

In order to assess the effects of Endow Iowa and the WPTC, I analyze contribution data reported on IRS 990 forms and compiled by the National Center for Charitable Statistics (NCCS). The NCCS Core files for 501(c)3 public charities include data on total revenue, contributions (inclusive of both grants and donations), membership dues, and program service revenue, as well as classifications from the National Taxonomy of Exempt Entities (NTEE).[12] The NCCS Core files include tax returns filed since the 1989 tax year. Since there is variation between the fiscal years of the nonprofits and the federal tax year in which the filings were received, I use datasets that begin with the 1990 fiscal year in all baseline estimates.

The primary outcome variable in this study is total contributions per capita. Total contributions include gifts, grants, and "membership dues"[13] which is taken directly from data in IRS form 990. For forms 990 submitted prior to 2008, this data is found in varying parts of the form, depending on the year: through 2005, line 1d in Part I; in 2006 and 2007, line 1e in Part 1; and from 2008 onward, line 1h in Part VIII. The 2008, 2009, and 2010 NCCS files include some returns filed using the "old" Form 990, which accounts for this inconsistency. The first page of the 2011 Form 990 appears in Figure 1.3. As the question of interest in this study is the overall economic impact

[12] The NTEE classifications consist of a letter followed by a two digit number. The letter classifications divide charities into broad categories such as "Education" (B) or "Mental Health and Crisis Intervention" (F). The two digit classifications further subdivides the charities into more specific categories such as "two-year colleges" (B41) or "substance abuse treatment" (F22).

[13] Membership dues are distinct from gifts in that the member receives services in return for their payment. As such, the payment is not tax deductible. A museum might provide free admission to members, who would only be able to claim payments made above and beyond the cost of membership as a charitable donations. Revenue from the museum gift shop, or single-day tickets would likely be counted as program service revenue.

of CTCs, total contribution is an ideal outcome variable. Tangentially, if one were to try to estimate the tax-price elasticity relative to CTCs, then the data from forms 990—and the related inability to differentiate between types of contributions and/or separate monetary from in-kind donations—would likely prove inadequate. By using this aggregate measure, I am able to incorporate substitution effects that may lead to a reduction in in-kind donations or private grants. Additionally, the aggregated contribution measure incorporates secondary effects on grants and membership that occur in response to changes in the level of private donations.

Although there is not a consistent measure of fundraising expenditure throughout the time series, I am able to derive a metric for each nonprofit in each year. Prior to 2008, the NCCS data includes total fundraising expenses taken directly from Part I line 15 of Form 990. I use this variable without adjustment. The 2008 version of Form 990 led to a critical adjustment in the NCCS data series. Data pulled from returns filed using the forms from 2008 onward do not include data for total fundraising.[14] Where total fundraising is unavailable, I construct an alternative measure by summing expenditures on professional fundraising services (Part IX Line 11e) and direct expenses from fundraising and gaming events (Part VIII Lines 8b and 9b).

I supplement the NCCS data with population estimates from the National Cancer Institute (SEER), inequality estimates from Frank (2009), and state-level personal income data from the Bureau of Economic Analysis (2014). Monetary data is inflated to 2012 dollars using the Bureau of Labor Statistics (BLS) Consumer Price Index - All Urban Consumers (CPI-U). Unemployment rate statistics from BLS are used to examine the robustness of counterfactuals.

Annual tax expenditures associated with Endow Iowa were reported in Gullikson and Tilkes' Endow Iowa Evaluation Study (2013). Information on grants distributed as part of the CEFP were provided by the Iowa Department of Revenue and are

[14] In 2008, 2009, and 2010, NCCS files include total fundraising for tax returns completed using old 990 forms.

available online at https://tax.iowa.gov/report/Distributions. Tax expenditures for the WPTC are available from annual versions of *The Revenue Impact of Arizona's Tax Expenditures* produced by the Arizona Department of Revenue's Office of Economic Research and Analysis.

1.3.2 Synthetic Control Methods

SCM were first described in Abadie and Gardeazabal (2003) and further developed in Abadie, Diamond, and Hainmueller (2010) and Abadie, Diamond, and Hainmueller (2015)— henceforth, ADH (2010) and ADH (2015). SCM is a novel approach that has been used in recent studies to examine economic liberalization (Billmeier and Nannicini, 2013), flat tax reforms (Adhikari and Alm, 2015), motion picture production tax incentives (Button, 2015), and the economic impacts of German reunification (Abadie et al., 2015), Norway's petroleum endowment (Mideksa, 2013), and two Italian earthquakes (Barone and Mocetti, 2014).

SCM create a "synthetic" counterfactual in which no policy change occurred. Mechanically, I create counterfactual versions of Arizona and Iowa as weighted averages of untreated states. The counterfactuals are created to mimic the trend in contribution levels to the targeted nonprofits before the tax credit policies were implemented. I then produce difference-in-difference (DID) estimates by comparing the Arizona and Iowa data with their synthetic counterfactuals.

Following ADH (2010) and ADH (2015), I create a synthetic state time series as a weighted combination of J untreated states in a "donor pool". (I describe the specific donor pools for Iowa and Arizona in greater detail in Section 1.3.3.) The weighting vector, $W = (w_2, ..., w_{J+1})$, is selected to minimize the distance between the treated state and its synthetic control during the pre-intervention period. Mechanically, W

solves the constrained minimization problem

$$W^* = \underset{W}{\operatorname{argmin}} \sqrt{(X_1 - X_0W)'V(X_1 - X_0W)} \tag{1.4}$$

$$s.t.$$

$$W'i = 1,\;\; w_j \geq 0,\;\; \text{for}\;\; j = (2, ..., J+1)$$

where X_1 is a vector of predictor variables for the treated state, X_0 is a matrix made up of vectors of predictor variables for each state in the donor pool, and V is a diagonal matrix that weights the relative importance of the predictor variables.

The primary outcome variable in my analysis is the natural log of per capita contributions. I use the per capita measure so that the population does not carry undue influence in determining the weights given to the states in the donor pool. As robustness checks, I mirror the analysis using per capita donations without the log transformation and with the natural log of total (rather than per capita) contributions.

The predictor variables in X_0 and X_1 serve to identify the combination of states that most closely match treated state time series to a counter-factual in the pre-intervention period. As such, they should be correlated with the outcome variable and not with the policy change. Note that the predictor variables do not need to have a causal relationship with contribution levels. If a predictor is highly correlated with an unobservable trait that affects contribution levels, it leads to a "better" synthetic control, even if there is no direct causal relationship.

I propose ten potential sets of predictor variables, listed in Table 1.3. The lists include four groups of potential predictors. The first is the variable of interest: contributions per capita or its natural log. Second, I include descriptors of the nonprofits of interest in the state: the per capita level of fundraising expenditures, and the per capita level of program revenue. These variables are associated with the underlying economic health of the nonprofits. The third group includes measures of the income distribution in each state: per capita income, the Gini coefficient, and the income share of the top one percent. We should expect giving to be higher in states with higher income levels. The relationship between inequality and charitable giving is not settled, but recent work by Payne and Smith (2015) suggests a positive relationship. Finally, the state population is included to account for differences in the nonprofit sector between large states and small states that may arise as a result of scale or agglomeration effects. With the exception of the inequality metrics, all predictor variables are log-transformed in the baseline estimate. Only the population variable is log-transformed in the robustness check based on per capita, rather than log per capita, contributions.

The period of time prior to policy implementation is used as a calibration period to select the best set of predictor variables. For each set, I generate synthetic controls for all states with complete data in the period from 1990 to 1994 and then estimate the average root mean square prediction error (RMSPE) in the following pre-intervention period: 1995 to 1998 in the case of Iowa and 1995 to 1997 in the case of Arizona.[15] RMSPE measures the degree to which the synthetic control fits the data series of the treated state and is defined as:

$$RMSPE = \left(\frac{1}{T_0} \sum_{t=1}^{T_0} (Y_t^{treated} - Y_t^{synth})^2 \right)^{\frac{1}{2}} \tag{1.5}$$

where $Y_t^{treated}$ and Y_t^{synth} are, respectively, the outcome variable and its synthetic counterfactual indexed by year t. The set of predictor variables that provides the best fit (lowest RMSPE) during the calibration period forms the vector X_1 and the matrix X_0 that are used to create the synthetic counterfactual.

After selecting the predictor variables, X_0, using the pre-intervention calibration

[15] The WPTC was implemented in 1998.

period, I construct a synthetic control using the period directly before intervention and estimate the difference between the treated state and its synthetic control over the post-intervention period. In the case of Iowa, the pre-treatment period runs from 1993 to 2002 and the post-treatment period runs from 2003 to 2012. For Arizona, the pre-treatment period runs from 1990 to 1997 and the post-treatment period runs from 1998 to 2012.

Once a synthetic control has been created, I estimate the difference-in-differences (DID) between the treated state and its counterfactual. The DID estimate represents the increase (or decrease) in the treated state relative to the increase (or decrease) in the synthetic control. In other words, the estimate for the causal effect of the policy is the change in the treated state minus the change that would have happened in the absence of the policy. Specifically, given any CTC policy p, define DD_p as:

$$DD_p = (\bar{Y}^{treated}_{post} - \bar{Y}^{synth}_{post}) - (\bar{Y}^{treated}_{pre} - \bar{Y}^{synth}_{pre}) \tag{1.6}$$

where $\bar{Y}^{treated}_{post}$ and $\bar{Y}^{synth}_{post}$ are the average annual (log) per capita contributions in the post-intervention period for the treated state and its synthetic control and $\bar{Y}^{treated}_{pre}$ and $\bar{Y}^{synth}_{pre}$ are the average annual (log) per capita contributions of the treated state and its synthetic control in the pre-intervention period. If the CTC is effective at increasing donations, the estimate of DD_p will be positive.

Additionally, I create alternative counterfactuals that I refer to as "Expected Iowa" and "Expected Arizona". The Expected state time series adds the tax expenditures associated with the CTC program. In Iowa, I also add the revenues distributed through the CEFP. The creation of these expected series provides two primary benefits. First, it makes it possible to estimate the benefits of the policies net of their costs. Second, it provides a test of Hypothesis 5 which proposes that the growth in contributions in Iowa should be greater than the value of the credits. In Iowa, I also construct a third counterfactual that includes the value of the grants provided through the CEFP but does not include the value of the credits distributed under Endow Iowa. This series provides an estimate of the gross effect of Endow Iowa alone, under the assumption that CEFP grants neither crowded in nor crowded out private contributions.[16]

Following ADH (2010), Billmeier and Nannicini (2013), and ADH (2015), I estimate p-values by running series of placebo experiments over the untreated states. The intuition is that by counting the number of times an estimate is larger for a placebo than it is following a policy change, we can calculate the probability that the estimate is a false positive. As such, I generate a synthetic control for each state in the donor pool (excluding the states with the highest and lowest average pre-treatment values) and calculate the corresponding DID estimator, DD_j. The p-value for DD_p is determined by its location within the distribution of estimators DD_j. If there are k placebo estimates larger than DD_p the p-value is calculated as $(k + 1)/(J + 1)$. Where DD_p is negative, I calculate the p-value as $((J - k) + 1)/(J + 1)$.

ADH (2015) propose the following ratio estimator for use in statistical inference:

$$Ratio_p = \frac{RMSPE_{post}}{RMSPE_{pre}}. \tag{1.7}$$

While the ratio estimator has no economic significance, it is used as an alternative statistic from which to generate p-values. Whereas the p-value associated with DD_p describes the likelihood that the average estimated difference between the treated group and its synthetic control could occur at random, the p-value associated with the $Ratio_p$ estimator describes the likelihood that the relative fit (or lack of fit) of the synthetic control could occur at random.

[16] This assumption is consistent with the model presented in Section 1.2.2.

1.3.3 Identification of Treatment Groups and Donors to the Controls

In Iowa, the treatment group consists of all community foundations in the state. The pre-qualified community foundations to whom donations are eligible for donor tax credits are a subset of a broader, nation-wide class of community foundations, which are identified under the National Taxonomy of Exempt Entities (NTEE) code **T31**. Comparing T31 community foundations in Iowa to those in other states, identifies "intent-to-treat" effect.

To construct Iowa's donor pool (the states from which the synthetic control will be constructed), I exclude all states that have, at any time since 1992, have given special tax treatment to planned gifts or donations to community foundations. This excludes Kansas, Kentucky, Michigan, Montana, North Dakota, and Nebraska. Next, I remove Arizona, which instituted a broad-based tax credit program (the WPTC). Hawaii and Utah are removed as well, because there is missing data. Finally, Wyoming and Delaware are removed from the baseline analysis in which per capita contributions are log-transformed because they have years of zero contributions. Both states are included in the untransformed robustness check. The remaining donor pool consists of 38 states and the District of Columbia (i.e. $J = 39$).

Figure 1.4a displays per capita contributions to community foundations in Iowa compared to those in the other 49 states and those in the donor pool. The vertical line indicates the implementation of the Endow Iowa program and thus separates the pre-treatment period from the post-treatment period. Iowa shows a large spike in per capita contribution in 2002.

The figure highlights an unusual and very large spike in Iowa's per capita contribution in 2002, most of which can be attributed to the Council Bluffs Community Betterment Foundation (CBCBF). In 2002, CBCBF received $75 million in contributions, representing the largest level of annual contributions to any Iowa community foundation in the dataset. This total is also six times higher than the maximum level of contributions Endow Iowa. This event appears to be an outlier, and hence, observations relating to the CBCBF are dropped from the baseline estimates. Furthermore, single year spikes in the level of the outcome variable make it difficult to find a strong synthetic control. Dropping the outlier allows for a stronger analysis and robustness tests show that it does not have a large effect on the results.[17] The resulting series of per capita contribution levels in Iowa is also demonstrated in Figure 1.4a in a different color.

The trends in Figure 1.4a show that Iowa had lower per capita giving to community foundations U.S. average in the pre-intervention period. Additionally, we can see the impact of the 2008 recession on giving—contributions drop significantly in 2009 for Iowa. The trend for the U.S. and the donor also shows a decrease in contributions, but not as significant as that in Iowa. It appears that Iowa's community foundations were hit harder by the recession than other states.

Table 1.4 displays the average per capita level of contributions to community foundations in Iowa (excluding CBCBF), the donor pool (from which the synthetic control is derived), and the United States as a whole. As demonstrated in the table, Iowa has more community foundations per person than the United States average. However, those foundations spend less on fundraising than their counterparts in the rest of the country. Looking at per captia income and the Gini coefficient, we can see that while slightly poorer, Iowa is slightly less unequal than both the donor pool and the national average. The control derived using SCM will empirically be more similar to Iowa than either the entirety of the donor group or the full sample of other states.

Since the rules that govern the WPTC are so broad, it is not possible to identify

[17] Alternative estimates based on aggregates including contributions to the CBCBF are presented along with other robustness checks. These results do not differ substantially from the main findings of the baseline analysis.

a single class of eligible nonprofits in Arizona. One option could be to examine the impact of the WPTC on an aggregate of the entire nonprofit sector. However, the value of the WPTC credits are equal to only 0.2 percent of all charitable receipts and therefore we are unlikely to be able to discern an impact at this level.[18]

Rather than identifying an intent-to-treat group, as in Iowa, in Arizona I focus on a group of "highly treated" nonprofits. In order to receive the tax credit, Arizona taxpayers must report which qualifying organization received their donation. A 2013 report by the Arizona Department of Revenue lists the 20 nonprofits that received the most donations qualifying for the tax credit in 2009 (Gene, 2013). Among these recipients, eight were part of national nonprofit networks with sister organizations in other states. Of these, six appeared annually in the NCCS data: Big Brothers Big Sisters, Boys and Girls Clubs, Goodwill Industries, Habitat for Humanity, St. Vincent De Paul Society, and United Way. I estimated the impact of the WPTC on an aggregate group made up of these six nonprofits. In 2009, credits received by donors to these nonprofits were equal to two percent of their charitable receipts. For comparison, the value of Endow Iowa credits is roughly three percent of total contributions to community foundations.

This highly treated group of six nonprofits is not a random subset of the treatment group. Rather, these six nonprofits represent organizations that were not only treated by the policy but were also, endogenously, recipients of larger-than-average levels of credit-eligible donations. The resulting analysis therefore estimates both a local average treatment effect among well-organized nationally-networked nonprofits and an upper-bound estimate of the average treatment effect. That is, this framework tests the hypothesis that charitable contributions to *some nonprofits* increased as a result of the WPTC.

[18] Using the methodology described in Section 1.3.2 and a state-wide nonprofit aggregate I find that, compared to a synthetic control, Arizona nonprofits received eight percent fewer contributions *after* the introduction of the WPTC.

The donor pool for the highly treated nonprofits in Arizona is formed from their counterparts in other states. Six states—Kansas, North Carolina, West Virginia, Michigan, and Missouri—are excluded due to CTC programs for which the six highly treated nonprofits were eligible or due to important policy changes during the period under study.[19]

Figure 1.4b displays per capita contributions to the highly treated nonprofits in Arizona compared with the contribution levels to their sister organizations in the other 49 states and DC and in the donor pool. The vertical line represents the introduction of the WPTC. The figure illustrates that Arizona saw an increase in contributions to the highly treated group in the period leading up to the introduction of the WPTC. Using SCM, a control group can be derived that accounts for this pre-existing trend.

Summary statistics for the highly treated nonprofits in Arizona and their sister organizations in the donor states and the United States as whole are displayed in Table 1.5. The table shows that per capita incomes in Arizona are below the national average. Notably, total fundraising spending in Arizona is, on a per capita basis, less than one seventh of the national average. The methodology described above addresses these differences in two ways. First, the DID estimator controls for differences in levels, relying instead on an assumption of identical trends pretreatment. Second, the synthetic Arizona is created specifically to match Arizona's pretreatment trends.

[19] I exclude Kansas due to its Community Service Tax Credit Program which provides credits for contributions to approved capital fundraising drives by nonprofits in community service, crime prevention, and healthcare access (Kansas Department of Commerce, 2014). Local affiliates of The United Way, Habitat for Humanity, Boys and Girls Clubs, and Big Brothers Big Sisters all qualified for credits under the program (Kansas Department of Commerce, 2014). North Carolina is removed because of a program that, until 2014, allowed non-itemizers to take a 7 percent credit on charitable donations in excess of 2 percent of their incomes (Revenue Research Division, 2013). Two states are removed due to programs with eligibility requirements that are similar to the WPTC. Virginia's Neighborhood Assistance Program and West Virginia's Neighborhood Investment Tax Credit provide credits for charitable gifts to nonprofits that meet certain criteria regarding their work with low income residents (in Virginia) or economically disadvantaged areas (in West Virginia) and are available to both individuals and businesses (Department of Social Services, 2014; West Virginia State Tax Department, 2012). Finally, I exclude Michigan and Missouri from the donor pool because each state had a variety of tax credits and a number of policy changes over the period of analysis.

1.4 Results

1.4.1 Iowa

The baseline estimates, displayed in Figure 1.5a, point to a large increase in contributions to community foundations in Iowa. The estimate of the net change in contributions for the causal impact of Iowa's policy (a combination of Endow Iowa Tax Credit and the CEFP) is a 63 percent increase in per capita contributions (a change of 0.49 in log levels). The estimate of the gross change in contributions suggests a 125 percent increase (a change of 0.81 in log levels). This finding represents more than \$45 million in additional philanthropic spending annually at a cost of less than \$6 million in lost tax revenue. Figure 1.5a displays the time series of log per capita contributions to community foundations in Iowa and its baseline synthetic control. As seen in the figure, contributions to community foundations in Iowa rose sharply beginning in the second year after the introduction of Endow Iowa. The level of giving predicted by the synthetic control, in contrast, shows little evidence of an increasing trend after 2000.

The synthetic Iowa in the baseline estimate is comprised of a weighted average of the eight states that lead to the best fit along a set of predictor variables, which are listed in line 9 of Table 1.3. Predictor list 9, which includes the pre-intervention average levels of all of the potential predictor variables along with the level of contributions in the first and last years of the pre-treatment period, was selected using the calibration process described in Section 1.3.2. The weighting matrix W used to create the synthetic control is displayed in Table 1.6. Table 1.6 shows that Vermont is given the greatest weight (26.7 percent) followed, respectively, by West Virginia, Ohio, Indiana, Maryland, Louisiana, Wisconsin, and Arkansas.

The results of the placebo experiments, displayed in Figure 1.5b, indicate that estimated increases in contributions of the magnitude found after introduction of Endow Iowa are unlikely to have occurred in the absence of a CTC policy. The solid black line indicates the gap in log per capita contributions to community foundations between Iowa and its synthetic control. The gray lines display the difference between the placebo states and their respective synthetic controls. Positive (negative) values indicate years in which contributions to Iowa community foundations or a placebo state were higher (lower) than to the synthetic control.

Contributions to community foundations are driven by large, one-time donations and, as such, can be erratic. Despite the volatility in the data, Iowa generally appears toward the top of the distribution during the post-intervention period. The p-values associated with the estimate of the gross change in contributions are estimated to be 0.03 using the distribution of DID estimates and 0.18 using the distribution of ratio estimators. The p-values associated with the estimate of the net change in contributions of additional giving beyond the cost of the program, are 0.13 and 0.18, respectively. Although we can reject the null hypothesis of no increase in contributions we cannot reject the hypothesis that the increase in contributions was no greater than the cost of the credits in terms of tax expenditures.

The most critical threat to a causal interpretation of these estimates is the possibility that the divergence between Iowa and its synthetic control was induced by some shock other than the introduction of Endow Iowa tax credits. Annual variation in contribution levels is too large to differentiate between Endow Iowa and the CEFP at the aggregate level. The SCM analysis presented here estimates the average treatment effect across the post-treatment period, and it is unable to differentiate between two closely related and closely timed policies. The firm-level estimates described in Section 1.5.4, however, find some evidence of different treatment effects. Additionally, any shock to the economy or market for charitable giving that occurred in Iowa or the synthetic control, but not the other, could cause contribution levels to diverge. It is not possible to fully rule out the possibility of all potential, unobservable shocks. However, we can rule out the possibility of any shocks that are highly correlated with observable trends.

To check for these observable trends, Iowa and its synthetic control are compared across state-wide economic and government spending indicators. The six measures that are used, which would likely pick up evidence of such a shock, are: (1) fundraising expenses, (2) unemployment, (3) per capita income, (4) inequality, (5) per capita state and municipal expenditures, and (6) per capita state and municipal revenue. All monetary metrics are measured in log form .[20] The results, illustrated in figure 1.6 show that the synthetic control generally follows the same trends as Iowa for every measure except for fundraising expenditure. Therefore, it is unlikely that there were notable, unobservable shocks that were closely correlated with the state-wide economy or government spending. However, as shown in the top left panel of Figure 1.6, the synthetic Iowa time series displays erratic levels of fundraising expenditures by community foundations that poorly fits the true Iowa time series. Notably, a finer analysis of the underlying data shows that a number of the donor states, especially Vermont, reported years in which no community foundation reported any fundraising expenditures. These zero values occurred in smaller states with few active community foundations. A zero value does not mean that the organizations did not solicit donations, but rather that they did not record expenditures as part of fundraising campaigns or events.

Because of it is possible that the lack of fundraising expenditures in some states points to fundamental differences between those community foundations and those in Iowa, and to control for the possibility that the erratic fundraising levels in the synthetic Iowa are symptoms of either unobserved shocks to the control or a vastly

[20] Synthetic control values displayed in Figure 1.6 are weighted averages of logged values rather than the log of weighted averages. This is consistent with the calculation of synthetic controls described in Section 1.3.2.

different environment that makes it a poor control, an alternative synthetic Iowa is created in which fundraising expenditure is the primary variable of interest. The synthetic control for fundraising expenditures by community foundations in Iowa is displayed in Figure 1.7. I then use the weights which optimized pre-treatment fit in fundraising to create an alternative time series for contribution levels. Using this new time series as a control for contributions in Iowa leads to an estimate of a 106 percent increase (an estimate of the gross change in contributions of 0.72 and an estimate of the net change in contributions of 0.36 in log levels) in contributions after the introduction of Endow Iowa. These results are only slightly smaller than those in the baseline. As such, shocks correlated with fundraising are unlikely to interfere with a finding of causality.

1.4.2 Arizona

The results from the baseline model show evidence of a small decline in contributions (9 percent, or -0.10 in log levels) to the group of highly treated nonprofits in Arizona after the introduction of the WPTC. Figure 1.8a displays the time series of per capita contributions to the group of nonprofit charities in Arizona and their synthetic control. On a per capita basis, WPTC tax expenditures grew from \$0.14 in its first year (1998) to \$3.33 in 2012. In 2009, the only year for which data on credit awards were available at the organizational level, donations to the six highly treated nonprofits accounted for 24 percent of all WPTC credits. The Expected Arizona time series assumes that this percentage—rather than the level of credits awarded—held constant from 1998 to 2012. Figure 1.8a shows that after the introduction of the tax credit, in more years than not, these six nonprofit received lower levels of contributions in Arizona than a synthetic counterfactual. The majority of that difference comes after 2004. Results of the placebo experiments are shown Figure 1.8b and produce a p-value of 0.35. An average decline of nine percent occurred in the placebo experiments about one third of the time, indicating that these results are not statistically significant.

The synthetic Arizona in the baseline estimate is comprised of a weighted average of eight states that lead to the best fit along a set of predictor variables. These predictor variables, listed in line 3 of Table 1.3, include the average of and the values for the first and last year in the pre-intervention period for all the potential predictor variables. In the baseline estimate, more than 60 percent of the variable weight (matrix V) is given to the average population in the pre-intervention period. As a robustness check, a new synthetic control is derived without using population as a predictor variable. The weighting matrix W used to create the synthetic control is displayed in Table 1.7. Georgia dominates in terms of weight, making up more than 49 percent of the control. Robustness checks in which each of the positively weighted states were (individually) excluded from the donor pool were also performed. The results of these robustness checks are discussed, along with others, in the following section.

As with Iowa, the data used in the WPTC analysis is investigated for evidence of some abnormality in Arizona or a donor state that may have biased the estimates. Figure 1.9 compares Arizona and its synthetic control across fundraising expenses, unemployment, per capita income, inequality, per capita state and municipal expenditures, and per capita state and municipal revenue, with all monetary metrics measured in log form.[21] There is some deviation in fundraising expenditures after the WPTC went into effect, but this result could be endogenous. The synthetic control tracks the true Arizona across all other measures closely, even during times of economic downturns.

[21] As was the case with Figure 1.6, Synthetic control values displayed in Figure 1.9 are weighted averages of logged values—rather than the log of weighted averages.

1.4.3 Robustness

The estimated effects of Endow Iowa and Arizona's WPTC are based on comparisons with a synthetic counterfactual. To check for robustness of the results, the treatment effects of both programs are re-estimated under alternative methodological choices and using alternative control groups. These approaches address two main concerns: bias(es) resulting from methodological choices and/or, as with any DID-based estimator, bias(es) resulting from the potential of an observable change in the control group(s).

The methodological choices described in Sections 1.3.2 and 1.3.3 do not appear to have affected the estimates. Results from robustness checks on Endow Iowa appear in Table 1.8. The removal of the apparent outlier, contributions to CBCBF, improved the precision of the estimates (i.e., a lower p-value associated with the ratio metric) but did little to affect the estimated treatment effect. The choice to convert monetary values into natural log form before estimation follows the existing literature. Yet, since SCM does not make the same normality assumptions as traditional regression-based analysis and because I estimate an average treatment effect rather than a pseudo-elasticity, this decision is arbitrary. SCM analysis without converting to natural log form suggests a $14.21 increase in contributions, and placebo estimates suggest a lower statistical likelihood of a false positive. Similarly, the analysis is iterated using levels of contributions at the aggregate level (i.e., not per capita), but still log-transforming the data. The results, in the last line of Table 8, estimate increases in contributions that are similar to the baseline.

The calibration procedure used to choose a combination of predictor variables proved not to be robust across choices of pre-treatment time period. Therefore, the results are re-examined using each of the remaining potential predictor variable lists shown in Table 1.3. The range of estimates of the treatment effect appears similar, regardless of which group of predictor variables is used. List 9, used in the baseline for Iowa, produces the largest estimate. The smallest estimate still suggests a 55 percent increase (0.44 in log levels) in contributions and a 26 percent (0.24 in log levels) increase beyond the costs of Endow Iowa and the CEFP.

I further examine whether the Iowa results might rely on any one state being included as a donor state by iterating the analysis and excluding a different individual donor state each time. The results appear both in Table 1.8 and Figure 1.10. In Figure 1.10, the back line represents the gap between log per capita contributions in Iowa and its baseline synthetic control. Each gray line represents the gap re-estimated with a donor state excluded. Removing Vermont or West Virginia from the donor pool reduces the estimated impact of Endow Iowa, while removing any of the other donor states has almost no effect. Repeating the analysis excluding Vermont or West Virginia, however, still results in an estimated (gross) increase in contributions of more than 50 percent. Estimated effects using non-log transformed per capita contributions were also re-estimated with the removal of each donor state, as well without the two states that had years of zero contributions (Wyoming and Delaware). Again, results were similar to those that appear in Table 1.8.

One might question whether geographically dissimilar states such as Maryland or Vermont can contribute appropriately to a counterfactual Iowa. For this reason, I perform a robustness check in which the donors are limited to neighboring states. The resulting synthetic control is comprised of Missouri (83 percent), Wisconsin (13 percent), and South Dakota (four percent) and provides the only negative point estimate. An alternative, albeit arbitrary, comparison of Iowa to the average across each of its neighboring states suggests an estimated (gross) increase in contributions of 36 percent.

Robustness checks for the WPTC in Arizona appear in Table 1.9. Notably, the baseline estimated decline in contributions is not robust across specifications. However, given the magnitude and the p-values associated with the baseline estimate, this result is not surprising. Estimates range from a 36 percent increase to a 21 percent decrease in contributions, but rarely is there a p-value below 0.2. Much of the negative result may be driven by the presence of Georgia in the donor pool, but even when Georgia is excluded, the estimates are smaller and the p-values are larger than those found in Iowa.

1.4.4 Secondary Hypotheses

I am unable to either reject Hypothesis 2 or its alternative. Hypothesis 2 predicts that the WPTC led to increased donations to untargeted nonprofits. Table 1.10 displays the SCM-based DID estimates of spillover effects which provide no evidence of an increase in contributions to nonprofits in Arizona that were unlikely to qualify under the WPTC. Nonprofits are defined as untargeted if they were not classified under NTEE codes J, L, O, P, or T, which would define them as focusing on the areas targeted by the WPTC—Employment, Housing and Shelter, Youth, Human Services, or Philanthropy. I also examine the possibility of spillover to affiliates of the group of "highly treated" nonprofits in neighboring states. No clear pattern emerges that would support or refute Hypothesis 2.

Tests of Hypothesis 3 also prove inconclusive. The theoretical model presented in Section 1.2.2 predicts that Endow Iowa should have induced substitution away from untargeted charities toward community foundations in Iowa. If this is the case, we should see a decline in contributions to other nonprofits, relative to their counterfactuals.

I examine two types of nonprofits that may be perceived as substitutes to community foundations in Iowa. The first category are nonprofits in Iowa that would not have qualified for the credit. including all other nonprofits, other nonprofits classified in the broad Public and Societal Benefit Sector (NTEE codes R, S, T, U, V, and W), and other nonprofits classified as belonging to either the community improvement and capacity building sub-sector (NTEE code S) or the philanthropy, volunteerism, and grant making sub-sector (NTEE code T). The results are mixed. In the case of the smallest and most similar comparison group, I find a decline in contributions, but one that is not statistically significant; it is replicated in 38 percent of placebo experiments. I find an increase in contributions to nonprofits in the broader Public and Societal Benefit sector and a decline in the category of all other nonprofits.

The second category of substitutes is community foundations in the states that border Iowa. Here, too, the results are mixed. A large increase in contributions to community foundations in South Dakota is statistically significant at the $\alpha = 0.1$ level, although the ratio test fails to meet that threshold. Note that three of the other five states that border South Dakota—Montana, Nebraska, and North Dakota—all enacted programs to promote planned giving and endowed charitable funds in the 1990s and 2000s. The increase in contribution levels in South Dakota could be the result of spillover effects from any or all of these programs. No clear pattern emerges from the results in Minnesota, Wisconsin, Illinois, and Missouri.

Empirical results are generally supportive of Hypotheses 4 and 5 but fail to reject the alternatives. Hypothesis 4 predicts that any increase in contributions to targeted nonprofits induced by the WPTC would be less than the value of the distributed credits. The baseline estimate points to a reduction in contributions to targeted nonprofits following implementation of the WPTC. While the baseline estimate of the net change in contributions (which subtracts the value of the credits from the change in contribution levels) in Table 1.9 is negative, it is not possible to reject the alternative hypothesis (that credits caused an increase in contributions that is positive even after subtracting the value of the credits from the change in contribution levels) at even the $\alpha = 0.1$ level of confidence. Hypothesis 5 predicts that Endow Iowa produced an increase in contributions that is greater than the associated value of the tax credits. The baseline point estimates of net change in contributions in Table 1.8 are positive and robust across a range of specifications. Again, however, the placebo tests suggest that the estimates are not strong enough to reject the alternative hypothesis.

1.5 Extensions and Inference on the Causal Mechanisms

1.5.1 Modest Growth in Per-Foundation Contribution Levels in Iowa

In order to understand the mechanism by which contributions increased in Iowa, it is necessary to estimate the effect of the policies on the number of foundations and the average effect of the policies at the organization level. This approach is analogous to studies of individual giving that differentiate between changes at the intensive and extensive margins. Here, rather than trying to decipher whether people are giving more or if there are more givers, I examine whether foundations are receiving more or if there are more foundations.

Panel DID methods are used to estimate the impact of the policies on average contributions per foundation. As discussed in Section 1.3.2, DID is used to estimate the increase in donations relative to the expected level in the absence of the policy. While SCM are ideal for case studies in which the outcome of interest is a state-wide aggregate, at the organization level, a fixed effects regression is more appropriate for analysis at the organizational level. The following form is used:

$$Y_{ist} = \alpha_s + \lambda_t + \delta^* Iowa^* Post + [X_{ist}\beta] + \epsilon_{ist} \tag{1.8}$$

where i indexes the organization and t indexes the year. In alternate specifications, s indexes treatment, state, or organization ($s = i$). The coefficient of interest, δ, is the DID estimator. As such, it is analogous to DD_p from equation 1.6. I take the natural log of all monetary variables, and convert them to per capita measures, before running the regression. The control vector, X_{ist}, includes state-level per capita income, population, Gini coefficient, and top one percent income share as well as charity-level fundraising expenditure and program revenue. Initially, I estimate the equation using a ten year balanced panel from 1998 to 2007.

If exogenous changes in Iowa's economy or in the behavior of community foundations both influenced the policy change and affect donations, estimates of the treatment effect will be biased without the inclusion of the control variables in X_{ist}. However, it is also possible that the policy change produced endogenous changes to the economy or to nonprofit behavior. If this is the case, the inclusion of controls X_{ist} would produce biased estimates of the treatment effect. Since it is unclear whether changes in economic conditions and nonprofit behavior are exogenous or endogenous, the regression is run both with and without controls.

The results are shown in columns 1 and 2 of Table 1.11 and suggest that contribution levels in Iowa increased by more than 25 percent following the passage of Endow Iowa. Columns 3 and 4 show that this finding is robust to the inclusion of state-level and firm-level controls. Columns 3 and 4 display estimates with and without fundraising, which almost certainly has an effect on contributions but is also likely to be endogenous to the treatment. The results are similar under both specifications.

Table 1.12 displays the sensitivity of the results to panel construction. I find a positive treatment effect when I iterate the analysis on both shorter and longer balanced panels. Moreover, the treatment effect remains statistically significant at the $p < 0.05$ level. However, the result does not hold when I run the same regression on the larger unbalanced panel. If the unbalanced panel is biased by attrition, then the estimates in Table 1.12 from the balanced panels are better estimates of the causal effect. With the exception of the longest, narrowest panel (column 3), estimates of the per-foundation treatment effect are still far below the synthetic control-based estimates of aggregate change. An increase in the number of community foundations would explain this discrepancy.

1.5.2 Growth in the Number of Community Foundations

To estimate the impact of the Endow Iowa tax credit on the number of foundations, I follow the synthetic control methodology described in Section 1.3.2, and replace per capita contributions with the number of community foundations per million people. The results are presented in Figure 1.11. Panel (a) shows that the number of community foundations is rising (on a per capita basis) in both Iowa and its synthetic control. The growth in Iowa outpaces its control, and the DID estimate is positive (0.25 in log levels, or 28 percent). However, the period with the most rapid growth in the number of community foundations appears to be before implementation of Endow Iowa. Panel (b) of Figure 1.11 displays the results of the placebo tests. The estimated DID p-value is 0.18.

SCM estimates and robustness checks for the change in number of community foundations are displayed in Table 1.13. The baseline estimate suggests a 28 percent increase in the number of community foundations. As above, I run robustness checks using the natural log of the outcome variable, including CBCBF in the Iowa data, using alternate predictor variable lists, and dropping each donor from the baseline estimate. The alternative specifications yield similar results.

Among tax-filers, the number of community foundations in Iowa grew from 35 in 2001 to 55 in 2012, representing 57 percent growth. Nationally, the number of community foundations grew 48 percent, while growth in donor states Indiana and Wisconsin was only 33 percent and 29 percent, respectively. The fact that the dataset only includes organizations that filed IRS 990 forms may actually understate the growth in the sector. A report by the Iowa Council of Foundations and Iowa Economic Development Authority (2014) states the number of community foundations in the state has grown to at least 130. Two types of community foundations would not appear in the dataset: those that received less than $50,000 in gross receipts (contributions plus other revenue), and affiliates of larger community foundations that filed 990 Forms jointly. Iowa's affiliate network is now extensive. Generally, these networks consist of newer, local (county) foundations and older, regional foundations (Iowa Council of Foundations and Iowa Gaming Association, 2013).

1.5.3 Endogenous Fundraising

It is possible that endogenous changes in fundraising could either increase or decrease the impact of a charitable tax credit. Figures 1.7 and 1.12 display the result of a synthetic control analysis for fundraising expenditure levels after implementation of Endow Iowa and the WPTC. There is no evidence of either a marked increase or decrease in fundraising levels relative to the synthetic control in either case. Arizona diverges from the synthetic control, but only during the period after the IRS Form 990 had changed.

Given that there appears to be an increase in the number of community foundations in Iowa and no change in aggregate fundraising levels, a decline in per-foundation fundraising expenditures should be expected. Table 1.14 displays estimates of the effect of Iowa's policy on fundraising by individual community foundations. I find a decline in per-foundation fundraising levels that is robust across a variety of specifications.[22]

The growth in the number of foundations, at least partially, explains how aggregate fundraising spending could remain steady while per-foundation fundraising

[22] Individual year effects λ_t account for the change in metrics for Fundraising Expenditure described in Section 1.3.1.

appeared to fall. The decline in fundraising expenditure at the foundation level could be the result of collaboration between foundations. Since the passage of Endow Iowa and the CEFP, community foundations have, in a sense, subdivided themselves. Donor response may be related to the combination of the new tax incentives and the opportunity to contribute to more localized public goods.

1.5.4 Disentangling Endow Iowa from the County Endowment Fund Grants

In order to disentangle the causal effects of Endow Iowa from the CEFP, I re-estimate equation 1.8 with unique treatment dummies for each. The results of these regressions, using a balanced panel from 1998 to 2007, appear in Table 1.15. The first two columns show an increase in contributions tied to the introduction of the Endow Iowa and a decline in contributions following the CEFP. After including a separate treatment variable for the CEFP, the estimate of increased contributions coincident with the introduction of Endow Iowa remains robust across the various balanced panels but, again, is not robust to the expansion to an unbalanced panel. The estimated decline in contributions coincident with the introduction of the CEFP is not robust to alternative specifications of balanced panels.

The estimates of changes in per capita fundraising expenditure in columns 3 and 4 of Table 1.15 further suggest that causal impacts are driven by Endow Iowa rather than the CEFP. Moreover, even after the inclusion of separate treatment effects, the estimated decline in fundraising is robust across the balanced and unbalanced panels.

If either policy had dynamic effects, this method is inconclusive. For example, it is possible that forward-looking residents anticipated the impacts of the County Endowment Fund program. Similarly, it is possible that the impact of Endow Iowa changed, for some reason other than interaction with the CEFP, concurrent with the initial disbursement of CEFP grants. This might happen if publicity surrounding the implementation of the tax credit produced a spike in giving in the first year that then tapered off. In either case, the differential effects of the policies cannot be uniquely identified.

1.6 Discussion

This paper provides the first rigorous, empirical analysis of the relationship between CTC policies and contributions to nonprofit organizations. Using SCM, I perform case studies of two contrasting CTC programs. In Iowa, I find as much as a 125 percent increase on per capita contributions to community foundations, the sector targeted by Endow Iowa and the CEFP. Increased contributions occurred as the number of community foundations rose. Foundations that existed at the time that Endow Iowa was implemented saw increased contributions, although the effect on average levels is less clear. I find no evidence that nonprofits in Arizona received higher levels of contributions in response to the WPTC. In the baseline estimate, point estimates suggest that Arizona's tax credit reduced contributions relative to a counterfactual.

It is important to differentiate the effects estimated in this paper from the tax-price elasticities estimated in the prior literature. Bakija and Heim (2011) and the extensive literature summarized by Peloza and Steel (2005) estimate the price elasticity of charitable giving using variation in the after-tax price of donations across time, donors, or both. In contrast, I estimate reduced form treatment effects in Iowa and Arizona. As such, the estimates in this paper include the effect of substitution *between* charities. Moreover, they include changes in contribution levels caused by behavioral responses to the policy changes such as changes in fundraising activities and the creation of new nonprofit entities.

The distinct results in Iowa and Arizona may be circumstantial, but there are four vital differences between the programs that could have caused Iowa's policy to outperform Arizona's. First, the structure of Arizona's credits—a 100 percent credit with a cap of $200 for individuals or $400 for couples—might have failed to provide appropriate incentives to donors. While more than $21 million in total credits were given to taxpayers in 2012 alone, the model presented in Section 1.2.2 suggests that most individuals would treat these funds as small $200 lump sum transfers. In contrast, the same model suggests that Iowa's program had the potential to induce substitution between charities. Second, Iowa sharply targeted its credit program to incentivize contributions to a specific sector. The growth in contributions coincided with—and may have produced—a growth in the number of foundations operating in the area that the policy targeted. The Iowa legislature may have been pointing nonprofit entrepreneurs to an area of (perceived) need. It is possible that a broad policy, such as Arizona's, would fail to have a similar effect. Third, both business and individuals are eligible for Endow Iowa credits, while only individuals qualify for the WPTC in Arizona. Fourth, Iowa followed their tax credit with the CEFP.

While this paper does not fully differentiate between the effects of Endow Iowa and the CEFP, the literature on grants to nonprofit organizations suggests that it would be unusual for a grant program to induce the large effects found in Iowa. Tinkelman (2010) summarizes 46 empirical studies and finds only six with estimated levels of additional giving greater than 50 percent. In contrast, I estimate additional giving of more than five times the amount granted through the CEFP. Where governments grants are found to crowd in additional giving, it is often proposed that they serve as a signal of quality. Indeed, Heutel (2014) finds larger levels of crowd in for younger charities, although those estimates would also suggest that the grants alone would not have caused contribution levels to more than double.

If the process by which states qualify organizations for CTCs provides additional, positive information to donors, CTCs might induce larger increases in contributions. Vesterlund (2003) provides a model in which early contributions provide signals of quality and produce higher contribution levels; lab experiments support the model's conclusions (Potters et al., 2005). Discussing universities, Payne (2001) explains that if donors possess imperfect information, government grants can serve to inform potential contributors of research quality and thus crowd in (increase) contributions. Qualifying for a CTC program may act as a similar signal. Heutel (2014) further posits that the signaling effect should be greater among charities about which less is known and finds (using nonprofit age as a proxy) evidence to support this hypothesis. Unlike traditional charities, community foundations are tasked not only with spending donors' money to produce public goods, but also with investing and managing the donors' money in order to maximize long-term impact. Insofar as this additional responsibility would make foundations less understood by donors than traditional charities (such as homeless shelters and food pantries), there will be a larger signaling effect produced by Endow Iowa and the CEFP than by the WPTC.

In Iowa, where the credit program appears to be a success, further research should consider second- and third-order effects. How are community foundations spending their additional income? Which nonprofits benefit from the growth in grant-making community foundations? Who benefits from their work? Additionally, it is worth further examining the relationship between the proliferation of community foundations and increased donations. Specifically, to what extent are donors drawn to more localized public goods?

The social welfare benefits of Endow Iowa are also of interest. Community foundations invest funds in endowments and distribute grants over time. As such, any substitution in philanthropic giving away from nonprofits that directly provide public services and into community foundations benefits those with longer time horizons and harms individuals whose time horizons are shorter. On the other hand, if donors substitute from spending on personal consumption to donations to community foundations that provide increased funding for public goods, social welfare will be improved as long as the public goods produced with these funds exceed the value of the public goods that the state could have produced with tax revenue lost to the credits. This study's examination of the nonprofit sector was unable to identify a substitution effect, but further research into individual expenditures may be able to do so.

These results represent the first systematic analysis of the effects of state-level charitable tax credits on nonprofits. Further research is needed, however, to discern the external validity of these results. SCM offer a focus on causal identification and quantitative rigor to case-study methodology and can be implemented to study a variety of other CTC programs. Finding appropriate identification strategies will be difficult, especially for programs that qualify projects rather than organizations. At this time, it isn't possible to fully identify which of the many differences between Endow and the WPTC drove the different results. As the literature on CTCs grows, research can turn to focusing on which features of the policies drive additional contributions and which do not.

Table 1.1: Summary of Selected Charitable Tax Credits

Credit	State	Effective	Ended	Issued (2012)	Qualifying Organizations	Pre-Qaulifying Required?	Project Specific?	Peronal/ Business	Percentage	Cap	Refund-able?	Carry-Over?	Reference
Education Tax Credit	AK	1987	–	$3.8 million	Nonprofit or public schools and colleges	No	No	Business	50%*	$5 million	No	No	(Alaska Department of Revenue, 2014) (Alaska Department of Revenue, 2015)
Working Poor Tax Credit	AZ	1998	–	$21.8 million	Varied	Yes	No	Personal	100%	$400/$800	No	Forward 5 years	(Gene, 2013) (Office of Economic Research and Analysis, 2014)
Neighborhood Assistance Tax Credit	CT	1982	–	$5 million	Varied	Yes	Yes	Business	60%**	$150,000	No	Back 2 years	(Office of Fiscal Analysis, 2012) (Conn. Gen. Stat. tit. 12 Ch. 228a §12-630aa-638)
Neighborhood Assistance Tax Credit	DE	2000	–	est. $200,000-$300,000	Varied	Yes	Yes	Business	50%	$100,000	No	Forward 5 years	(Division of Revenue. State of Delaware, 1999) (Del. Code tit. 30 Ch. 11 §. 2001-2007) (Department of Finance, 2011)
Endow Iowa Tax Credit	IA	2003	–	$5.8 million	Community Foundations	Yes	No	Both	25%	$300,000	No	Forward 5 years	(Gullickson and Tilkes, 2013)
Community Service Tax Credit Program	KS	1994	–	$4.1 million	Community Service, Crime Prevention, and Health Care Nonprofits	Yes	Yes	Both	50%***	$250,000 per Organization	Yes	No	(Kansas Department of Commerce, 2014)
Endow Kentucky	KY	2011	–	$200,000	Community Foundations	Yes****	No	Both	20%	$10,000	No	Forward 5 years	(Governor's Office for Economic Analysis, Office of State Budget Directory, 2011) (Ky. Rev. Stat. §141.438)
Donations to Resource and Referral Agencies	LA	2008	–	$218,539	Private Agencies with contracts through the Department of Social Services	Yes	No	Business	100%	$5,000	Yes	No	(Louisiana Department of Revenue, 2013) (Louisiana Department of Revenue, 2015)
Homeless Shelter / Food Bank Credit	MI	1992	2011	$20.0 million (2011)	Homeless Shelters and Food Banks			Both	50%	$100/$200 (Individuals) $5,000 (Businesses)			(Tax Analysis Division, Office of Revenue and Tax Analysis, 2014)
Community Foundation / Education Credit	MI	1989	2011	$3.8 million (2011)	Community and Education Foundations			Both	50%	$100/$200 (Individuals) $5,000 (Businesses)	No	No	(Tax Analysis Division, Office of Revenue and Tax Analysis, 2014)
Youth Opportunities Program	MO	1996	–	$3.8 million	Varied	Yes	Yes	Both	50%	$200,000	No	Forward 5 years	(Missouri Department of Economic Development, 2015) (Missouri Department of Economic Development, 2013)
Food Pantry Tax Credit	MO	2007	2011	$793,794 (2010)	Food Pantries	No	No	Both	50%	$2,500	No	Forward 3 years	(Oversight Division, 2011)
Qualified Endowment Credit	NE	2006	2009	$150,000 (2008)	Any 501(c)3 with an endowment	No	No	Both	15%*****	$5,000	No	No	(Nebraska Department of Revenue Research Division, 2008) (Nebraska Department of Revenue, 2010)
Donations to Biomedical Research Institutes	OK	2005	–	$514,000	Medical Research Insitutes	No	No	Both	50%	$1,000	No	Forward 4 years	(The Tax Policy Division of The Oklahoma Tax Commission, 2012) (Okla. Admin. Code §710:50-15-113)

* Alaska Education Credit is available for up to 50% of annual contributions up to $100,000, 100% of the next $200,000, and 50% of annual contributions beyond $300,000.
** Connecticut provides a 100% credit for energy conservation projects and construction or rehabilitation of low-income housing units.
*** The Kansas Community Service Program 70% credits for contributions in rural areas.
**** Endow Kentucky requires preliminary authorization be requested by the donor rather than the grantee organization.
***** Nebraska's Qualified Endowment Credit provided a 15% credit for individuals, S corporations, partnerships and limited liability companies and a 10% credit for C corporations.

Table 1.2: Predicted Effects of Tax Credits

Preferences	Endow Iowa 1 (Targeted)	Endow Iowa 2 (Untargeted)	WPTC (Arizona) 1 (Targeted)	WPTC (Arizona) 2 (Untargeted)
Linear Substitutes, $1 \succ 2$	$\uparrow$ 33%	$=$	$\uparrow$ \$200	$=$
Linear Substitutes, $2 \succ 1$	$\Uparrow$	$\downarrow$	$\uparrow$ \$200	$=$
Linear Substitutes, $2 \succ\succ 1$	$=$	$=$	$\uparrow$ \$200	$=$
Cobb-Douglas Preferences	$\uparrow$ 33%	$=$	$\uparrow<$ \$200	$\uparrow<$ \$200
CES, Gross Substitutes, $1 \succ 2$	$\Uparrow$	$\downarrow$	$\uparrow<$ \$200	$\uparrow<$ \$200
CES, Gross Substitutes, $2 \succ 1$	$\Uparrow$	$\downarrow$	$\uparrow<$ \$200	$\uparrow<$ \$200
CES, Gross Complements, $1 \succ 2$	$\uparrow$	$\uparrow$	$\uparrow<$ \$200	$\uparrow<$ \$200
CES, Gross Complements, $2 \succ 1$	$\uparrow$	$\uparrow$	$\uparrow<$ \$200	$\uparrow<$ \$200
Perfect Complements	$\uparrow$ 14%	$\uparrow$ 14%	$\uparrow<$ \$200	$\uparrow<$ \$200

Notes: The table displays the change in a donor's contribution level under the assumptions of the model presented in section 2.2. $\Uparrow$ is used to signify an increase in donations that includes substitution from untargeted to targeted nonprofits. $1 \succ 2$ ($2 \succ 1$) denotes that the targeted (untargeted) charity would receive a larger donation than the untargeted (targeted) charity in the absence of a credit. $2 \succ\succ 1$ implies that the the preference for the untargeted charity is strong enough that it receives all donations whether or not there is a credit.

Table 1.3: Choices of Predictor Variables

	Predictor Variables
1	Annual Per Capita Contributions
2	Average Per Capita Contributions, Program Revenue, and Fundraising Expenditures; Average Per Capita Income, Population, Gini, and Top 1 Percent Share
3	Average, and the values for the first and last years in the pre-intervention period of: Per Capita Contributions, Per Capita Program Revenue, Per Capita Fundraising Expenditures, Per Capita Income, Population, Gini, and Top 1 Percent Share
4	Annual Per Capita Contributions; Average Program Revenue and Fundraising Expenditures; Average Per Capita Income, Population, Gini, and Top 1 Percent Share
5	Annual Per Capita Contributions; Average, and the values for the first and last years in the pre-intervention period of: Per Capita Program Revenue, Per Capita Fundraising Expenditures, Per Capita Income, Population, Gini, and Top 1 Percent Share
6	Average Per Capita Program Revenue and Fundraising Expenditures; Average Per Capita Income, Population, Gini, and Top 1 Percent Share
7	Average, and the values for the first and last years in the pre-intervention period of: Per Capita Program Revenue, Per Capita Fundraising Expenditures, Per Capita Income, Population, Gini, and Top 1 Percent Share
8	Average Per Capita Contributions; Average, and the values for the first and last years in the pre-intervention period of: Per Capita Program Revenue, Per Capita Fundraising Expenditures, Per Capita Income, Population, Gini, and Top 1 Percent Share
9	Average, and the values for the first and last years in the pre-intervention period of Per Capita Contributions; Average Per Capita Fundraising Expenditures and Program Revenue; Average Per Capita Income, Population, Gini, and Top 1 Percent Share
10	Average, and the values for the first and last years in the pre-intervention period of Per Capita Contributions Program Revenue, and Fundraising Expenditures

Notes: In all cases, I use the natural log of population. When creating a synthetic counterfactual time series of the natural log of per capita contributions, as in the baseline model, per capita fundraising expenditure, program revenue, and income are also log-transformed.

Table 1.4: State-level Summary Statistics: Community Foundations

	Iowa		Donor Pool		U.S.	
	Mean	sd	Mean	sd	Mean	sd
Total Per Capita Contributions	18.02	11.69	19.87	28.67	18.72	26.09
Per Capita Program Revenue	0.30	0.64	2.30	12.94	1.85	11.37
Per Capita Fundraising Expenditures	0.10	0.08	0.29	0.52	0.26	0.47
Community Foundations per million people	12.17	3.80	4.27	2.68	5.17	4.16
Population (Millions)	2.95	0.07	6.53	6.82	5.66	6.30
Per Capita Income	30,598	7,678	32,423	9,255	31,974	8,965
Gini Coefficient	0.55	0.01	0.59	0.04	0.59	0.04
Top 1 Percent Income Share	0.14	0.03	0.17	0.04	0.17	0.04

Notes: Standard deviations appear in the columns labeled “sd”. All statistics are average levels for the years 1993 to 2012 and measured in 2012 dollars. Contributions, Program Revenue, and Fundraising expenditures for Council Bluffs Community Betterment Foundation are not included in the summary statistics.

Table 1.5: State-level Summary Statistics: Six National Nonprofit Charities

	Arizona		Donor Pool		U.S.	
	Mean	sd	Mean	sd	Mean	sd
Total Per Capita Contributions	22.12	5.41	23.47	22.31	23.11	20.89
Per Capita Program Revenue	7.50	5.68	7.02	8.04	6.87	7.61
Per Capita Fundraising Expenditures	1.22	0.55	9.98	269.34	8.76	250.18
Population (Millions)	5.24	0.93	5.52	6.59	5.56	6.20
Per Capita Income	26,944	6,909	30,80	9,693	30,285	9,465
Gini Coefficient	0.58	0.02	0.59	0.04	0.58	0.04
Top 1 Percent Income Share	0.16	0.03	0.16	0.04	0.16	0.04

Notes: Standard deviations appear in the columns labeled “sd”. All statistics are average levels for the years 1989 to 2012 and measured in 2012 dollars. Contribution, Revenue, and Expenditure statistics are state-wide aggregates of all affiliates of Big Brothers Big Sisters, Boys and Girls Clubs, Goodwill Industries, Habitat for Humanity, St. Vincent De Paul Society and United Way.

Table 1.6: Weights for Donor States, Iowa Community Foundations

State	Weight
Arkanas	0.027
Indiana	0.121
Louisiana	0.059
Maryland	0.108
Ohio	0.179
Vermont	0.267
Wisconsin	0.029
West Virginia	0.210

Notes: Table displays donor weights (matrix W). States not listed were either given zero weight or excluded from the donor pool. Weights create a synthetic aggregate of community foundations at the state-level.

Table 1.7: Weights for Donor States, Arizona Affiliates of Six National Nonprofit Charities

State	Weight
Alaska	0.016
Florida	0.005
Georgia	0.493
Idaho	0.038
Oklahoma	0.092
South Carolina	0.039
Texas	0.039
Utah	0.277

Notes: Table displays donor weights (matrix W). States not listed were either given zero weight or excluded from the donor pool. Weights create a synthetic aggregate of Big Brothers Big Sisters, Boys and Girls Clubs, Goodwill Industries, Habitat for Humanity, St. Vincent De Paul Society and United Way at the state-level

Table 1.8: Estimated Policy Impact on Contributions to Foundations in Iowa

	Gross Change			Net Change		
	treatment effect	p-values DID	p-values Ratio	treatment effect	p-values DID	p-values Ratio
Baseline Estimate	0.81	0.03	0.18	0.49	0.13	0.18
CBCBF not dropped	0.80	0.03	0.21	0.45	0.16	0.21
Predictor Variable List 1	0.50	0.05	0.13	0.25	0.24	0.13
Predictor Variable List 2	0.75	0.08	0.11	0.47	0.16	0.11
Predictor Variable List 3	0.80	0.03	0.08	0.50	0.08	0.08
Predictor Variable List 4	0.50	0.08	0.16	0.25	0.21	0.16
Predictor Variable List 5	0.57	0.08	0.16	0.32	0.24	0.16
Predictor Variable List 6	0.71	0.08	0.11	0.34	0.32	0.11
Predictor Variable List 7	0.73	0.05	0.11	0.38	0.13	0.11
Predictor Variable List 8	0.75	0.05	0.08	0.45	0.13	0.08
Predictor Variable List 10	0.44	0.16	0.26	0.24	0.26	0.26
Arkansas Excluded	0.80	0.05	0.22	0.48	0.16	0.22
Indiana Excluded	0.80	0.03	0.19	0.47	0.11	0.19
Louisiana Excluded	0.74	0.03	0.16	0.46	0.14	0.16
Maryland Excluded	0.80	0.03	0.16	0.47	0.14	0.16
Ohio Excluded	0.81	0.03	0.19	0.50	0.14	0.19
Vermont Excluded	0.48	0.08	0.41	0.28	0.19	0.41
Wisconsin Excluded	0.77	0.03	0.19	0.47	0.14	0.19
West Virginia Excluded	0.47	0.14	0.76	0.27	0.16	0.76
Weights for Fundraising	0.72	0.13	0.68	0.36	0.35	0.95
Only Neighboring States	-0.03	0.50	0.44	0.00	0.00	0.00
Excluding Population	0.91	0.03	0.21	0.55	0.11	0.21
Versus Regional Average	0.31	–	–	0.18	–	–
Versus National Average	0.35	–	–	0.15	–	–
Not Log Transformed	14.21	0.10	0.05	10.70	0.10	0.08
Not Per Capita	0.85	0.08	0.18	0.51	0.16	0.24

Notes: Treatment effect is a DID estimate of change in log per capita contributions after intervention. The "Net Change" estimate adds the tax expenditures associated with Endow Iowa and the grants distributed by the Community Endowment Program to the counterfactual before calculating the DID estimate. P-values are displayed for both the DID estimate and the ratio estimator proposed in ADH (2015). Baseline Estimate follows the methodology described in section 1.3.2. Estimates for untransformed per capita contributions follow similar methodology with an expanded donor pool that includes Wyoming and Delaware (which were dropped from the baseline due to years of zero contributions) and predictor set 2. Estimates of log, not per capita, contributions use predictor set 9. Unless otherwise specified, the remaining estimates use the same set of predictor variables as the baseline, but create alternative synthetic controls. The p-values displayed are based on the null hypothesis that Endow Iowa increased contributions.

Table 1.9: Estimated Policy Impact on Contributions to Six National Nonprofit Charities in Arizona

	Gross Change			Net Change		
	treatment	p-values		treatment	p-values	
	effect	DID	Ratio	effect	DID	Ratio
Baseline Estimate	-0.10	0.35	0.37	-0.11	0.33	0.37
Predictor Variable List 1	0.14	0.30	0.60	0.13	0.35	0.60
Predictor Variable List 2	0.16	0.35	0.30	0.15	0.37	0.30
Predictor Variable List 4	0.14	0.30	0.60	0.13	0.35	0.60
Predictor Variable List 5	0.14	0.30	0.60	0.13	0.35	0.60
Predictor Variable List 6	0.18	0.26	0.19	0.16	0.30	0.19
Predictor Variable List 7	0.02	0.49	0.26	0.01	0.49	0.26
Predictor Variable List 8	-0.13	0.33	0.28	-0.14	0.33	0.28
Predictor Variable List 9	0.00	0.51	0.37	-0.01	0.51	0.37
Predictor Variable List 10	-0.24	0.12	0.40	-0.25	0.12	0.40
Including 1989 Data	0.31	0.14	0.65	0.29	0.14	0.65
Excluding Population	0.14	0.30	0.60	0.13	0.35	0.60
Alaska Excluded	-0.17	0.29	0.36	-0.18	0.26	0.36
Florida Excluded	-0.11	0.29	0.38	-0.12	0.29	0.38
Georgia Excluded	0.23	0.12	0.36	0.22	0.14	0.38
Idaho Excluded	-0.08	0.36	0.36	-0.10	0.31	0.36
Oklahoma Excluded	-0.08	0.38	0.36	-0.09	0.36	0.36
South Carolina Excluded	-0.10	0.36	0.38	-0.12	0.31	0.38
Texas Excluded	-0.14	0.29	0.38	-0.15	0.26	0.38
Utah Excluded	-0.13	0.31	0.40	-0.14	0.29	0.38
Only Neighboring States	0.09	0.60	0.40	0.07	0.60	0.40
Versus Regional Average	0.10	–	–	0.08		–
Versus National Average	0.16	–	–	0.15		–
Not Log Transformed	0.97	0.49	0.58	0.97	0.49	0.58
Not Per Capita	0.17	0.19	0.60	0.40	0.09	0.30

Notes: Treatment effect is a DID estimate of change in log per capita contributions after intervention. The "Net Change" estimate adds the value of credits awarded (for donations to these six nonprofits) under the Working Poor Tax Credit to the counterfactual before calculating the DID estimate. The comparison is made between a state-level aggregate of Big Brothers Big Sisters, Boys and Girls Clubs, Goodwill Industries, Habitat for Humanity, St. Vincent De Paul Society and United Way and its synthetic control. P-values are displayed for both the DID estimate and the ratio estimator proposed in ADH (2015). Baseline Estimate follows the methodology described in section 1.3.2. Estimates for untransformed per capita contributions follow similar methodology with an expanded donor pool that includes Wyoming and Delaware (which were dropped from the baseline due to years of zero contributions) and predictor set 9. Estimates of log, not per capita, contributions use predictor set 9. Unless otherwise specified, the remaining estimates use the same set of predictor variables as the baseline, but create alternative synthetic controls. Where estimates are positive (negative)The p-values displayed are based on the null hypothesis that WPTC increased (decreased) contributions.

Table 1.10: Estimated Spillover Effects from Endow Iowa and the WPTC

	treatment effect	p-values DID	p-values Ratio
Estimates of Spillover from Endow Iowa			
Similar Nonprofits in Iowa	-0.11	0.39	0.58
Public/Societal Benefit Nonprofits in Iowa	0.09	0.21	0.21
Untargeted Nonprofits in Iowa	-0.06	0.29	0.53
Community Foundations in Nebraska	0.35	0.15	0.21
Community Foundations in South Dakota	0.63	0.05	0.26
Community Foundations in Minnesota	-0.33	0.24	0.32
Community Foundations in Wisconsin	0.00	0.47	0.39
Community Foundations in Illinois	0.09	0.39	0.97
Community Foundations in Missouri	0.61	0.08	0.92
Estimates of Spillover from WPTC			
Untargered Nonprofits in Arizona	-0.05	0.44	0.16
Affiliates in California	0.40	0.07	0.95
Affiliates in Nevada	-0.13	0.28	0.60
Affiliates in Utah	-0.25	0.12	0.09
Affiliates in Colorado	0.06	0.42	0.91
Affiliates in New Mexico	-0.03	0.42	0.81

Notes: Treatment effect is a synthetic control-based difference-in-difference estimate of change in log per capita contributions after intervention. The category "Similar Nonprofits in Iowa" includes (Form 990 filing) nonprofit organizations, other than NTEE code T31 community foundations, classified as belonging to either the community improvement and capacity building sub-sector (NTEE code S) or the philanthropy, volunteerism, and grant making sub-sector (NTEE code T). "Public/Social Benefit Nonprofits in Iowa" includes nonprofits, other than community foundations, classified under the broad Public and Societal Benefit Sector (NTEE codes R, S, T, U, V, and W). "Untargeted Nonprofits in Iowa" includes all nonprofits other than community foundations. The remaining spillover estimates for Iowa refer to changes in contributions to community foundations in neighboring states. "Untargeted Nonprofits in Arizona" includes all (Form 990 filing) nonprofit organizations that are not classified under NTEE codes J, L, O, P, and T, which would classify them as focusing on areas targeted by the WPTC—Employment, Housing and Shelter, Youth, Human Services, and Philanthropy. Additional spillover estimates in Arizona refer to the change in contributions at affiliates of Big Brothers Big Sisters, Boys and Girls Clubs, Goodwill Industries, Habitat for Humanity, St. Vincent De Paul Society and the United Way in neighboring states. Ratio is the ADH (2015) metric described in equation 1.7. Baseline Estimate follows the methodology described in sections 1.3.2.

Table 1.11: Change in Contributions Per Community Foundation (10 Year Balanced Panel)

	(1)	(2)	(3)	(4)
Treatment	0.268**	0.268**	0.253*	0.287*
	(0.085)	(0.090)	(0.119)	(0.121)
Iowa	-1.439***			
	(0.111)			
Per Capita Income			3.165	3.272
			(1.578)	(1.721)
Populaton			0.612	0.706
			(1.536)	(1.623)
Gini Coefficient			0.027	0.496
			(1.734)	(1.868)
Top 1				
			(1.989)	(2.079)
Fundraising Expenditure				0.022**
				(0.008)
Program Revenue				-0.006
				(0.005)
Year FE	yes	yes	yes	yes
Foundation FE	yes	yes	yes	yes
Observations	2650	2650	2650	2650

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Notes: Contributions, Fundraising Expenditure, Program Revenue, Per Capita Income, and Population are natural log transformed. All dollars are real, 2012 base year.Standard Errors (in parentheses) are clustered at the state level.

Table 1.12: Change in Contributions Per Community Foundation (Sensitivity)

		Balanced		Unbalanced
	1998-2007	2000-2005	1993-2012	
Treatment	0.287*	0.358***	0.729***	0.059
	(0.121)	(0.082)	(0.143)	(0.059)
Per Capita Income	3.272	1.894	2.429	2.118*
	(1.721)	(1.865)	(1.588)	(1.017)
Populaton	0.706	0.023	0.156	0.603*
	(1.623)	(0.119)	(1.174)	(0.277)
Gini Coefficient	0.496	2.543	8.612*	-1.449
	(1.868)	(2.724)	(3.164)	(1.833)
Top 1% Share	-1.954	-2.724	-8.229**	1.100
	(2.079)	(2.365)	(2.806)	(1.477)
Fundraising Expenditure	0.022**	0.019**	0.016**	0.016***
	(0.008)	(0.006)	(0.006)	(0.003)
Program Revenue	-0.006	0.002	-0.012	-0.001
	(0.005)	(0.007)	(0.009)	(0.003)
Year FE	yes	yes	yes	yes
Foundation FE	yes	yes	yes	yes
Observations	2650	2712	1700	13269

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Notes: Contributions, Fundraising Expenditure, Program Revenue, Per Capita Income, and Population are natural log transformed. All dollars are real, 2012 base year.Standard Errors (in parentheses) are clustered at the state level.

Table 1.13: Estimated Policy Impact on the Number of Community Foundations in Iowa

	treatment effect	p-values DID	p-values Ratio
Baseline Estimate	0.25	0.18	0.47
CBCBF not dropped	0.25	0.18	0.50
Predictor Variable List 1	0.25	0.18	0.47
Predictor Variable List 2	0.34	0.11	0.08
Predictor Variable List 3	0.33	0.13	0.05
Predictor Variable List 4	0.25	0.18	0.47
Predictor Variable List 5	0.25	0.18	0.47
Predictor Variable List 6	0.24	0.16	0.26
Predictor Variable List 7	0.23	0.18	0.29
Predictor Variable List 8	0.34	0.13	0.05
Predictor Variable List 10	0.30	0.16	0.11
Indiana Excluded from Donor Pool	0.26	0.16	0.54
Oklahoma Excluded	0.24	0.16	0.49
Wisconsin Excluded	0.25	0.19	0.51
Restriction to Neighboring States	0.25	0.25	0.25
Excluding Population as a Predictor	0.30	0.08	0.13
Not Log Transformed	3.62	0.03	0.08
Not Per Capita	0.08	0.32	0.82

Notes: Treatment Effect is a difference-in-difference estimate of change in (log) number of foundations per million residents or (log) per capita fundraising expenditures. Post/Pre ratio is the ADH (2014) metric described in equation 1.7. Baseline Estimate follows the methodology described in section 1.3.2. Estimates for untransformed measures were made by separately calibrating the list of predictor variables. Unless otherwise specified, the remaining estimates use the same set of predictor variables as the baseline, but create alternative synthetic controls. The p-values displayed are based on the null hypothesis that Endow Iowa increased either number of foundations.

Table 1.14: Change in Fundraising Expenditures Per Community Foundation

		Balanced		Unbalanced
	1998-2007	2000-2005	1993-2012	
Treatment	-1.710*	-1.102*	-1.167	-0.713*
	2.370	(0.449)	(0.991)	(0.330)
Per Capita Income	(0.766)	-7.588	-12.945	0.227
	(9.119)	(11.825)	(9.455)	(3.600)
Populaton	5.186	-0.801	3.486	-0.516
	(4.954)	(0.785)	(7.602)	(0.645)
Gini Coefficient	-28.118	-25.646	-32.689	-5.684
	(27.578)	(14.687)	(42.215)	(9.828)
Top 1% Share	28.636	32.085*	6.908	8.186
	(24.724)	(15.566)	(34.065)	(9.004)
Year FE	yes	yes	yes	yes
Foundation FE	yes	yes	yes	yes
Observations	2460	2712	1700	13269

* $p< 0.05$ ** $p< 0.01$ *** $p< 0.001$

Notes: Contributions, Fundraising Expenditure, Program Revenue, Per Capita Income, and Population are natural log transformed. All dollars are real, 2012 base year.Standard Errors (in parentheses) are clustered at the state level.

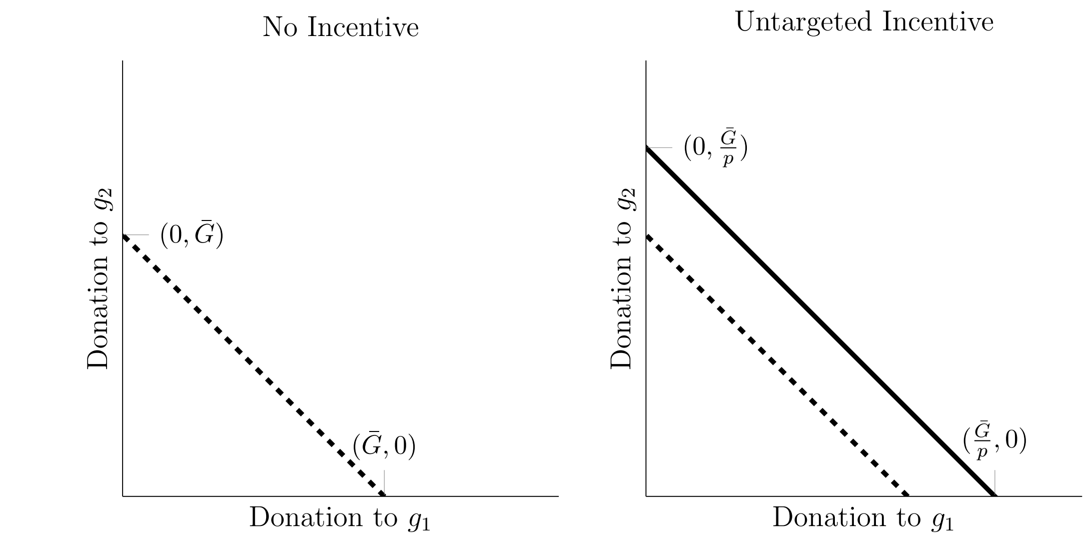
Table 1.15: Differential Impacts of Endow Iowa and Community Endowment Fund Program (1998-2007 Balanced Panel)

	Contributions		Fundraising	
	(1)	(2)	(3)	(4)
Endow Iowa	0.339**	0.378**	-1.721**	-1.941***
	(0.114)	(0.113)	(0.517)	(0.526)
Community Endowment Fund	-0.146**	-0.155**	0.452	0.524
	(0.048)	(0.050)	(0.290)	(0.314)
Per Capita Income	3.162	3.269	-3.341	-5.710
	(1.584)	(1.728)	(10.186)	(10.655)
Populaton	0.586	0.679	-1.512	-2.593
	(1.531)	(1.619)	(10.920)	(10.777)
Gini Coefficient	-0.039	0.427	-14.250	-16.005
	(1.739)	(1.873)	(17.417)	(17.954)
Top 1% Share	-1.191	-1.898	27.570	29.402
	(1.994)	(2.085)	(20.241)	(20.556)
Fundraising Expenditure		0.022**		
		(0.008)		
Program Revenue		-0.006		0.078
		(0.005)		(0.041)
Contributions				0.625**
				(0.213)
Year FE	yes	yes	yes	yes
Foundation FE	yes	yes	yes	yes
Observations	2650	2650	2650	2650

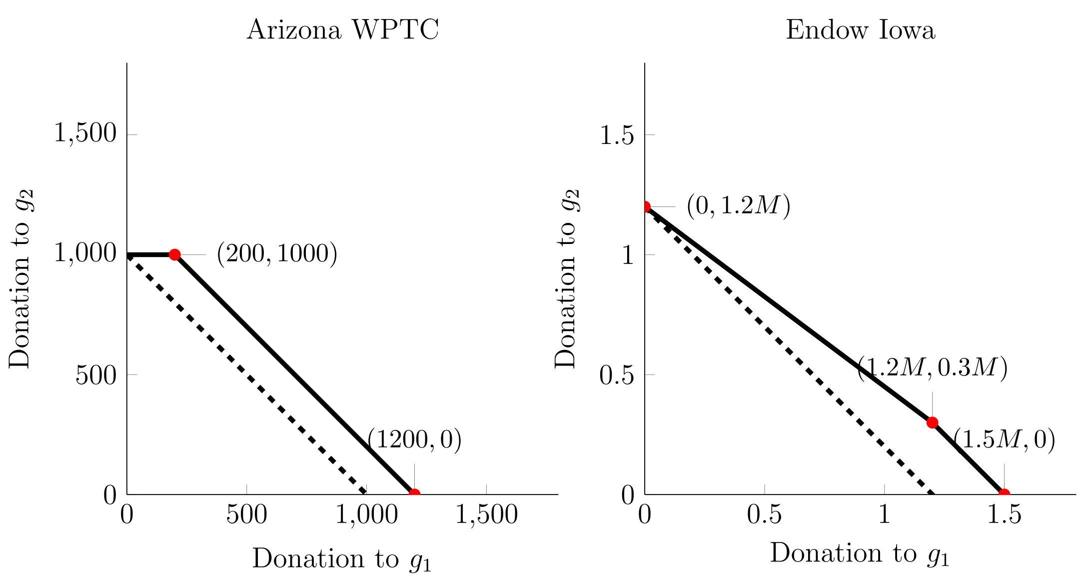
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Notes: Contributions, Fundraising Expenditure, Program Revenue, Per Capita Income, and Population are natural log transformed. All dollars are real, 2012 base year.Standard Errors (in parentheses) are clustered at the state level.

Figure 1.1: Budget Constraints of Donors in Arizona and Iowa

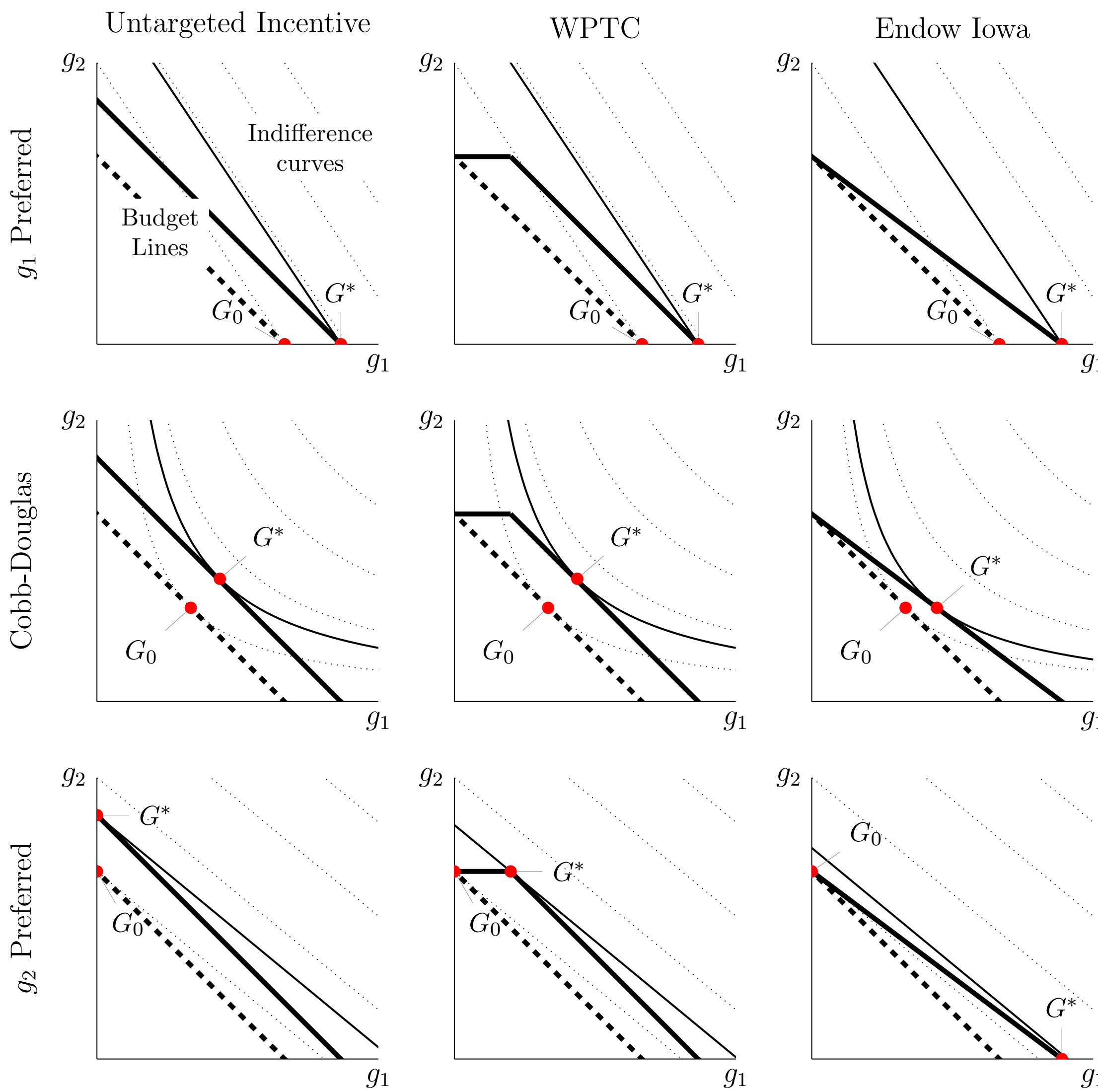


(a) Budget constraints before (dashed) and after (solid)the implementation of an untargeted tax incentive for charitable giving. The budget line shifts outward in proportion to the size of the incentive. The slope of the line is unchanged.



(b) Budget constraints before (dashed) and after (solid)the implementation of the CTC programs in Arizona and Iowa. Kinks in the budget constraint occur when the total subsidy reaches a statutory cap. Note that the budget constraints are not on the same scale. The graphs represent a donor qualifying for the WPTC while donating \$1,000 after taxes and a donor qualifying for Endow Iowa while donating \$1.2 million after taxes. The disparate budgets were chosen to make the kinks in the budget constraints visible in both graphs. In Figure 1.2, the budget constraints are on the same scale.

Figure 1.2: Consumer Responses to Changes in the Tax Treatment of Charitable Giving



A model of donation with targeted tax incentives. The figure displays the budget constraints before (dashed) and after (solid)the implementation of an untargeted incentive and CTC programs in Arizona (WPTC) and Iowa (Endow Iowa). Indifference curves in the top row assume the charities are substitutes but the targeted charity is preferred. Indifference curves in the middle row assume that donors see charities as neither substitutes nor complements. Indifference curves in the bottom row assume that the charities are substitutes, that the untargeted charity is preferred in the absence of a subsidy, and that the targeted charity is preferred when it is subsidized.

Figure 1.3: IRS Form 990

Form **990** — Department of the Treasury, Internal Revenue Service

Return of Organization Exempt From Income Tax

Under section 501(c), 527, or 4947(a)(1) of the Internal Revenue Code (except black lung benefit trust or private foundation)

▶ The organization may have to use a copy of this return to satisfy state reporting requirements.

OMB No. 1545-0047

2011

Open to Public Inspection

A For the 2011 calendar year, or tax year beginning , 2011, and ending , 20

B Check if applicable:
- ☐ Address change
- ☐ Name change
- ☐ Initial return
- ☐ Terminated
- ☐ Amended return
- ☐ Application pending

C Name of organization

Doing Business As

Number and street (or P.O. box if mail is not delivered to street address) | Room/suite

City or town, state or country, and ZIP + 4

D Employer identification number

E Telephone number

G Gross receipts $

F Name and address of principal officer:

H(a) Is this a group return for affiliates? ☐ Yes ☐ No

H(b) Are all affiliates included? ☐ Yes ☐ No
If "No," attach a list. (see instructions)

I Tax-exempt status: ☐ 501(c)(3) ☐ 501(c) ( ) ◀ (insert no.) ☐ 4947(a)(1) or ☐ 527

J Website: ▶

H(c) Group exemption number ▶

K Form of organization: ☐ Corporation ☐ Trust ☐ Association ☐ Other ▶ **L** Year of formation: **M** State of legal domicile:

Part I Summary

Activities & Governance

1	Briefly describe the organization's mission or most significant activities:			
2	Check this box ▶ ☐ if the organization discontinued its operations or disposed of more than 25% of its net assets.			
3	Number of voting members of the governing body (Part VI, line 1a)	3		
4	Number of independent voting members of the governing body (Part VI, line 1b)	4		
5	Total number of individuals employed in calendar year 2011 (Part V, line 2a)	5		
6	Total number of volunteers (estimate if necessary)	6		
7a	Total unrelated business revenue from Part VIII, column (C), line 12	7a		
b	Net unrelated business taxable income from Form 990-T, line 34	7b		

			Prior Year	Current Year
Revenue	8	Contributions and grants (Part VIII, line 1h)		
	9	Program service revenue (Part VIII, line 2g)		
	10	Investment income (Part VIII, column (A), lines 3, 4, and 7d)		
	11	Other revenue (Part VIII, column (A), lines 5, 6d, 8c, 9c, 10c, and 11e)		
	12	Total revenue—add lines 8 through 11 (must equal Part VIII, column (A), line 12)		
Expenses	13	Grants and similar amounts paid (Part IX, column (A), lines 1–3)		
	14	Benefits paid to or for members (Part IX, column (A), line 4)		
	15	Salaries, other compensation, employee benefits (Part IX, column (A), lines 5–10)		
	16a	Professional fundraising fees (Part IX, column (A), line 11e)		
	b	Total fundraising expenses (Part IX, column (D), line 25) ▶		
	17	Other expenses (Part IX, column (A), lines 11a–11d, 11f–24e)		
	18	Total expenses. Add lines 13–17 (must equal Part IX, column (A), line 25)		
	19	Revenue less expenses. Subtract line 18 from line 12		

			Beginning of Current Year	End of Year
Net Assets or Fund Balances	20	Total assets (Part X, line 16)		
	21	Total liabilities (Part X, line 26)		
	22	Net assets or fund balances. Subtract line 21 from line 20		

Part II Signature Block

Under penalties of perjury, I declare that I have examined this return, including accompanying schedules and statements, and to the best of my knowledge and belief, it is true, correct, and complete. Declaration of preparer (other than officer) is based on all information of which preparer has any knowledge.

Sign Here

Signature of officer | Date

Type or print name and title

Paid Preparer Use Only

Print/Type preparer's name | Preparer's signature | Date | Check ☐ if self-employed | PTIN

Firm's name ▶ | Firm's EIN ▶

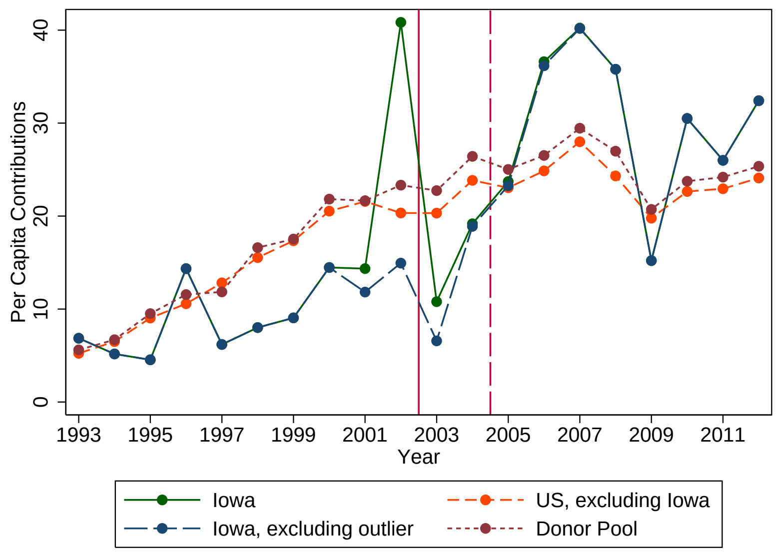
Firm's address ▶ | Phone no.

May the IRS discuss this return with the preparer shown above? (see instructions) ☐ Yes ☐ No

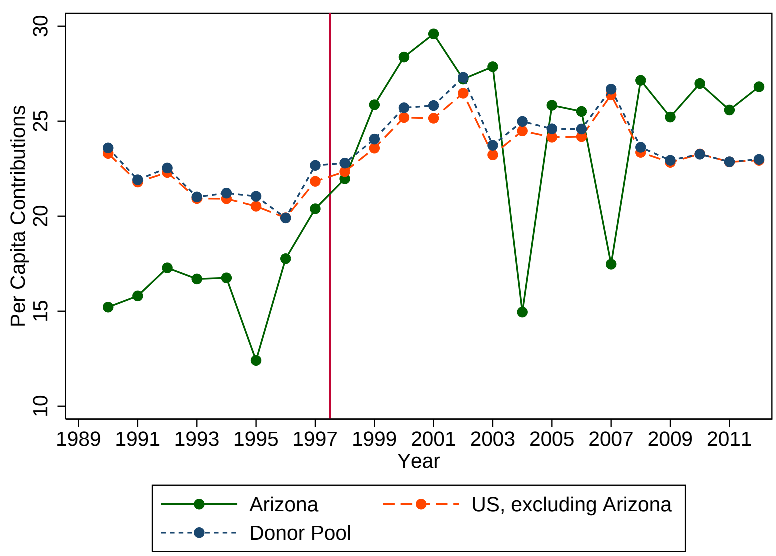
For Paperwork Reduction Act Notice, see the separate instructions. Cat. No. 11282Y Form **990** (2011)

Page 1 of a sample IRS Form 990 downloaded from the IRS website, http://www.irs.gov/pub/irs-prior/f990ez–2011.pdf . Lines 8-12 document sources of revenue. Contributions appear on line 8.

Figure 1.4: Per Capita Contributions to Treated Nonprofits and their Donor Pools.



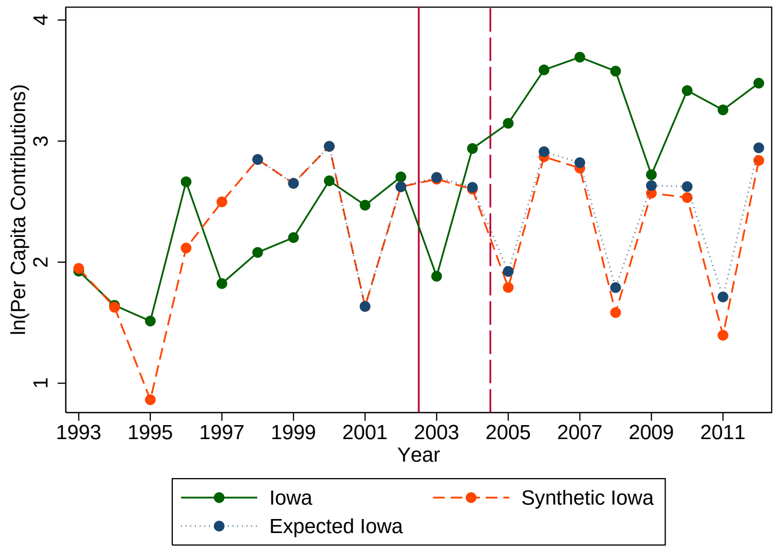
(a) Per Capita Contributions to Community Foundations in Iowa, the United States, and the Donor Pool. 1993–2012



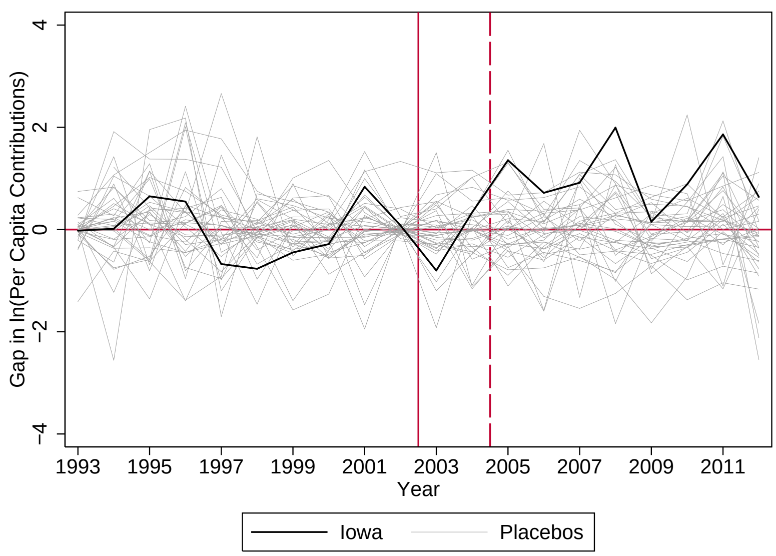
(b) Aggregate Per Capita Contributions to Six National Nonprofits in Arizona, the United States, and the donor pool, 1990–2012

Figures represent total per capita contributions reported by 501(c)3 nonprofit charities on IRS Form 990. Iowa, excluding outlier, removes Council Bluffs Community Betterment Foundation. The Donor Pool for Iowa consists of 39 untreated states. The Donor Pool for Arizona consists of 44 untreated states. Solid vertical lines represent the introduction of Endow Iowa and the WPTC. The dashed vertical line in panel (a) represents the first distribution of County Endowment Funds.

Figure 1.5: Synthetic Control Analysis: Endow Iowa Tax Credit



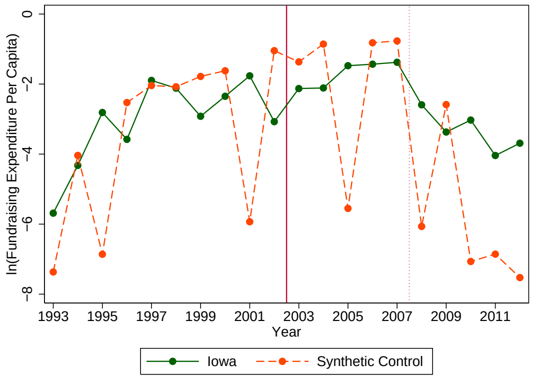
(a) Log Per Capita Contributions to Community Foundations: Iowa and Synthetic Controls



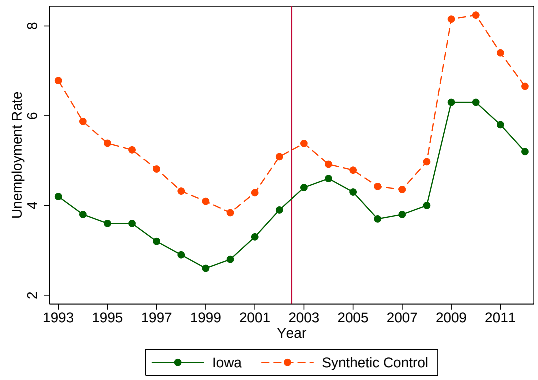
(b) Difference in the Log Per Capita Contributions relative to Synthetic Control: Iowa and Placebos

The Iowa time series in panel (a) represents total per capita contributions reported by community foundations on IRS Form 990. Synthetic Iowa is derived from a donor pool or 39 untreated states. Expected Iowa adds the costs associated with Endow Iowa and the County Endowment Fund program to the synthetic counterfactual. In panel (b) the black solid line displays the difference between Iowa and its synthetic control. Gray lines represent placebo tests. In both panels, the solid vertical line represents the introduction of the Endow Iowa Tax Credit. In panel (a) the dashed vertical line represents the first distribution of County Endowment Funds.

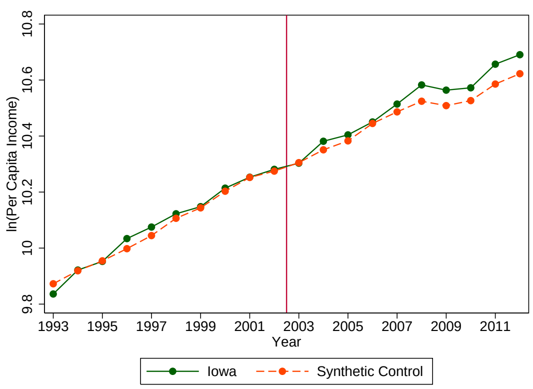
Figure 1.6: Balance between Iowa and its Synthetic Control: Trends across Six Measures



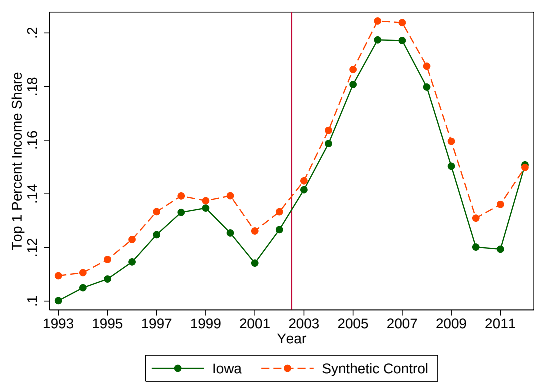
(a) Log Per Capita Fundraising Expenditures



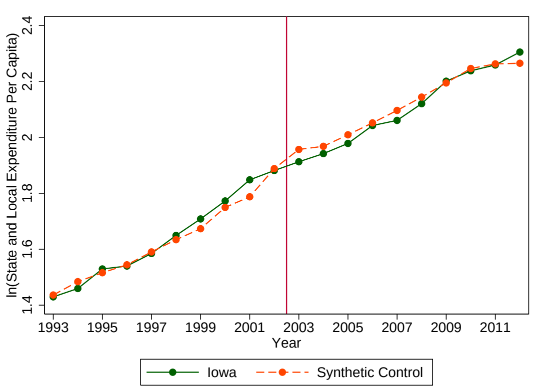
(b) State-wide Unemployment Rate



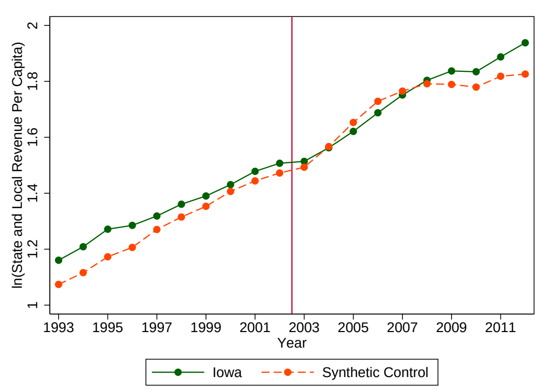
(c) Log Per Capita Income



(d) Inequality: Share of Income to the Top 1 percent



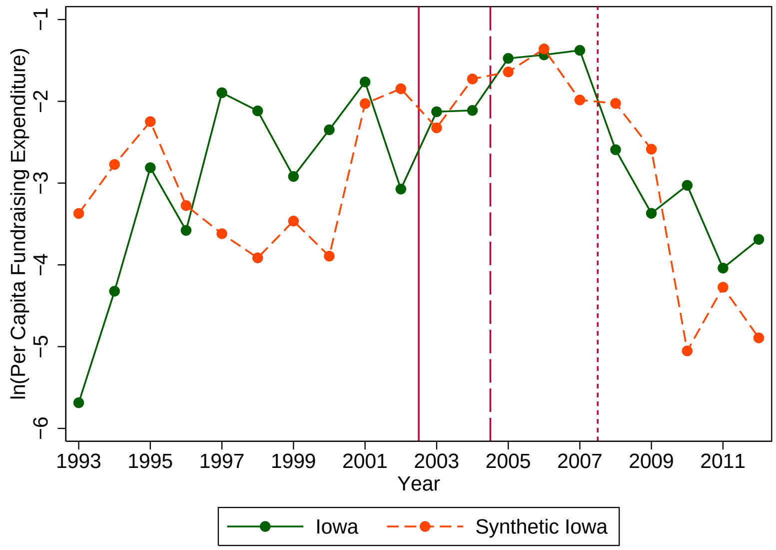
(e) Log State and Municipal Expenditures



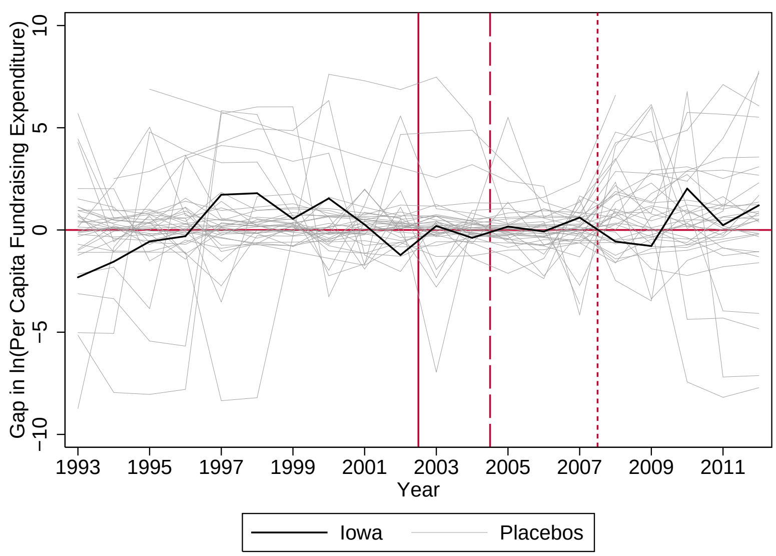
(f) Log State and Municipal Revenues

Note: The figures compare trends in Iowa to those of its baseline synthetic control (the synthetic control estimated to fit the trend in per capita contributions in the pre-intervention period shown in figure 1.5). Trends in fundraising, income, and inequality were used to fit the synthetic control. The solid vertical line separates pre- and post-intervention periods. The dotted line in panel (a) marks the change break in the data series between alternative fundraising expenditure aggregates.

Figure 1.7: Synthetic Control Analysis: Endow Iowa Tax Credit (Fundraising)



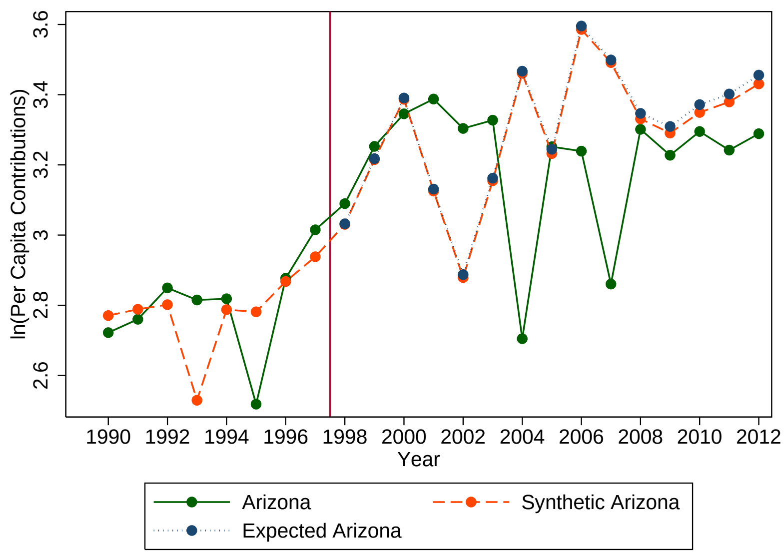
(a) Log Per Capita Fundraising Expenditures by Community Foundations: Iowa and Synthetic Controls



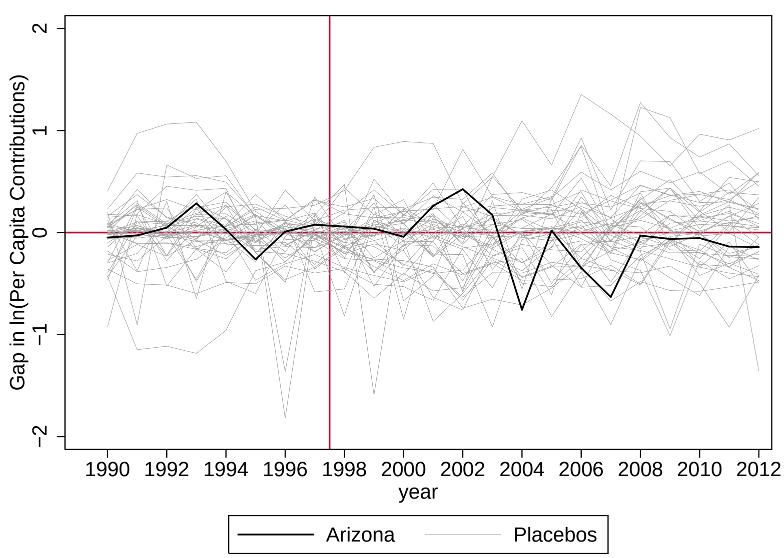
(b) Difference in the Log Per Capita Fundraising Expenditures relative to Synthetic Control: Iowa and Placebos

The Iowa time series in panel (a) represents total per capita fundraising expenditures reported by community foundations on IRS Form 990. Synthetic Iowa is derived from a donor pool of 39 untreated states. In panel (b), the black line displays the difference between Iowa and expected Iowa. Gray lines represent placebo tests. In both panels, the vertical lines represent the introduction of Endow Iowa Tax (solid) and the change in how fundraising expenditure is reported on IRS Form 990 (dashed).

Figure 1.8: Synthetic Control Analysis: The WPTC (Arizona)



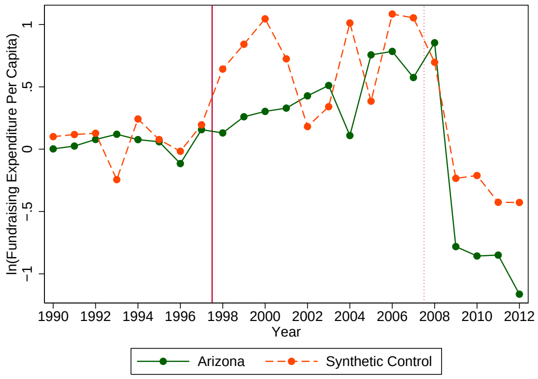
(a) Log Per Capita Contributions to Six National Nonprofits: Arizona and Synthetic Controls



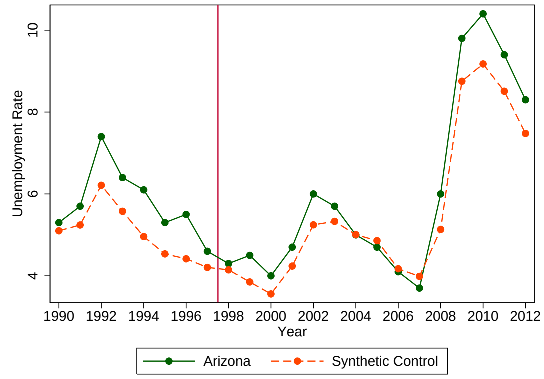
(b) Difference in the Log Per Capita Contributions relative to Synthetic Control: Arizona and Placebos

The Arizona time series in panel (a) represents total per capita contributions reported by six highly treated national nonprofit charities on IRS Form 990. Synthetic Arizona is derived from a donor pool of 43 untreated states. Expected Arizona adds the estimated tax expenditure associated with credits received by donors to the six nonprofits to the synthetic counterfactual. In panel (b), the black line displays the difference between Arizona and expected Arizona. Gray lines represent placebo tests. In both panels, the vertical lines represent the introduction of the WPTC.

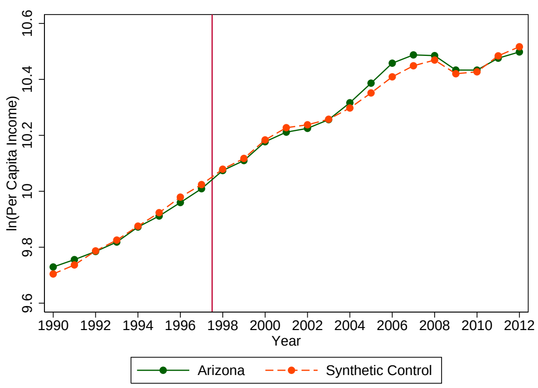
Figure 1.9: Balance between Arizona and its Synthetic Control: Trends across Six Measures



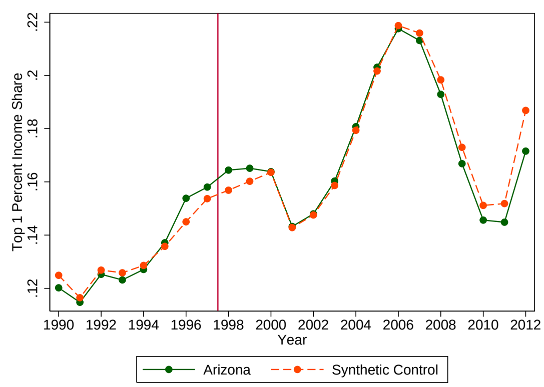
(a) Per Capita Fundraising Expenditures



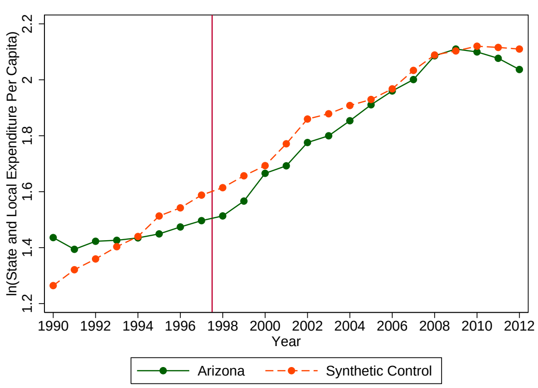
(b) State-wide Unemployment Rate



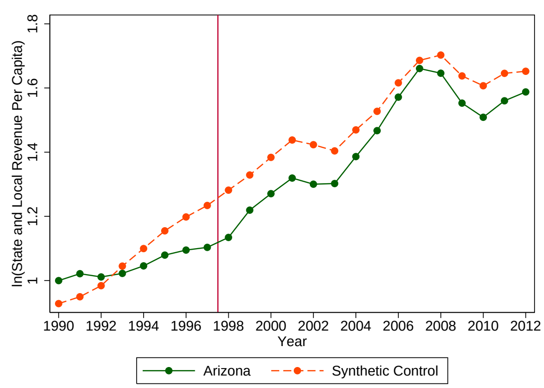
(c) Per Capita Income



(d) Inequality: Share of Income to the Top 1 percent



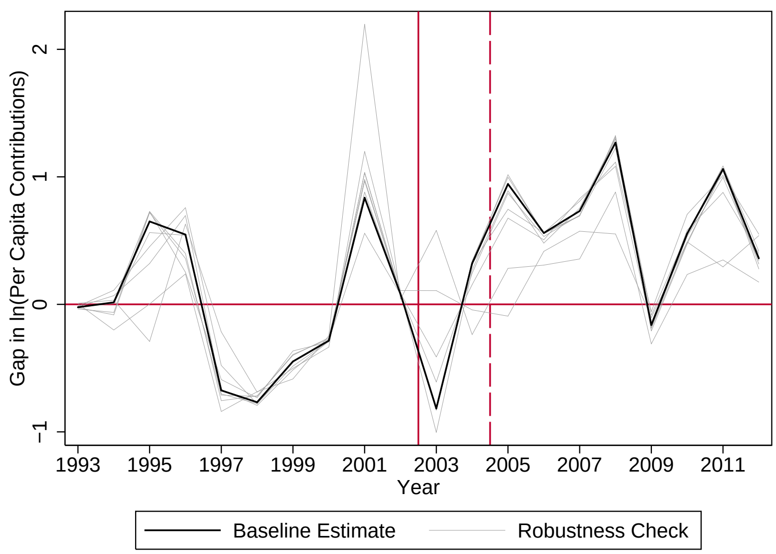
(e) State and Municipal Expenditures



(f) State and Municipal Revenues

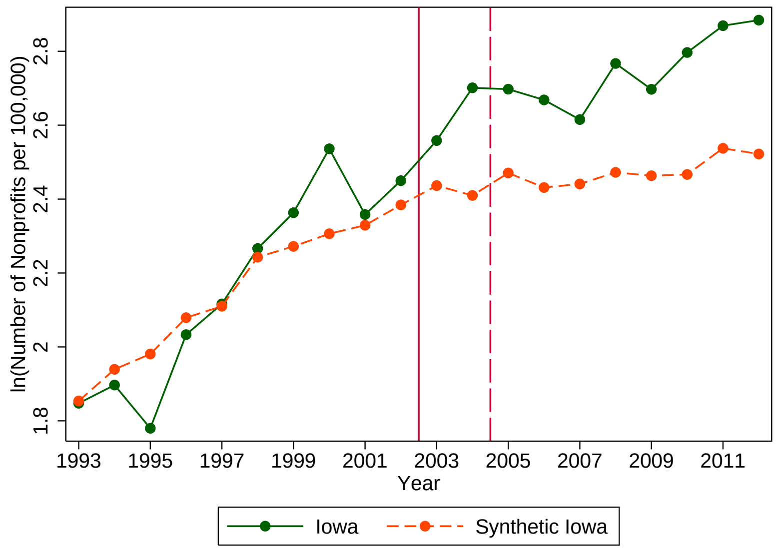
Note: The figures compare trends in Arizona to those of its baseline synthetic control (the synthetic control estimated to fit the trend in Per Capita Contributions in the pre-intervention period shown in figure 1.8). Trends in fundraising, income, and inequality were used to fit the synthetic control. The solid vertical line separates pre- and post-intervention periods. The dotted line in panel (a) marks the change break in the data series between alternative fundraising expenditure aggregates.

Figure 1.10: Leave-One-Out Robustness Check: Endow Iowa

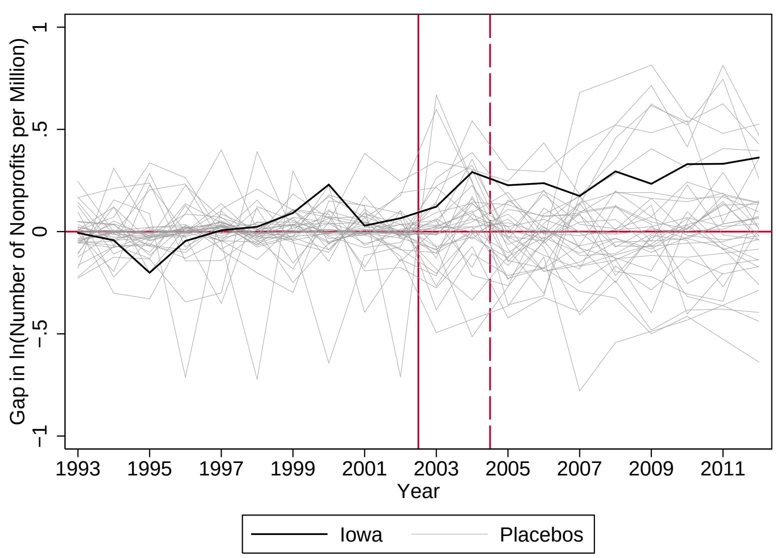


The black line displays the difference between Iowa and expected Iowa. Gray lines represent the difference calculated from alternative synthetic controls, in which one of the donor states from the baseline estimate is excluded. The vertical line represents the introduction of the Endow Iowa.

Figure 1.11: Synthetic Control Analysis: Endow Iowa Tax Credit (Community Foundations)



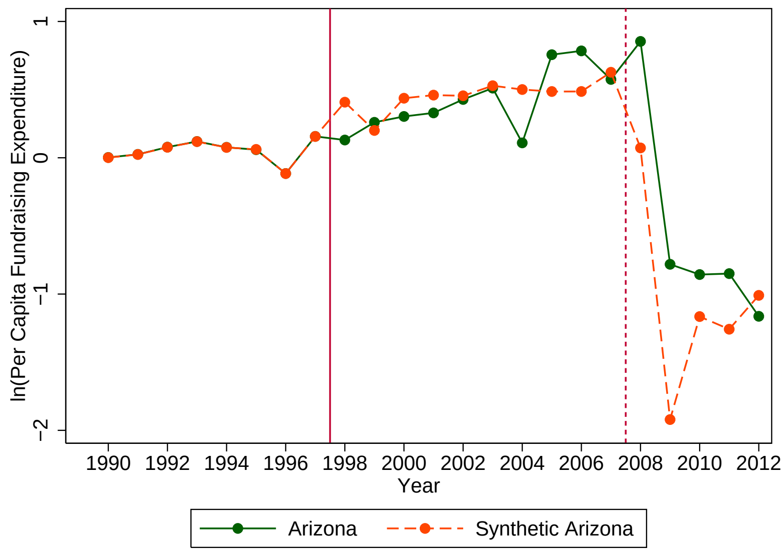
(a) Log of Community Foundations Per Capita: Iowa and Synthetic Controls



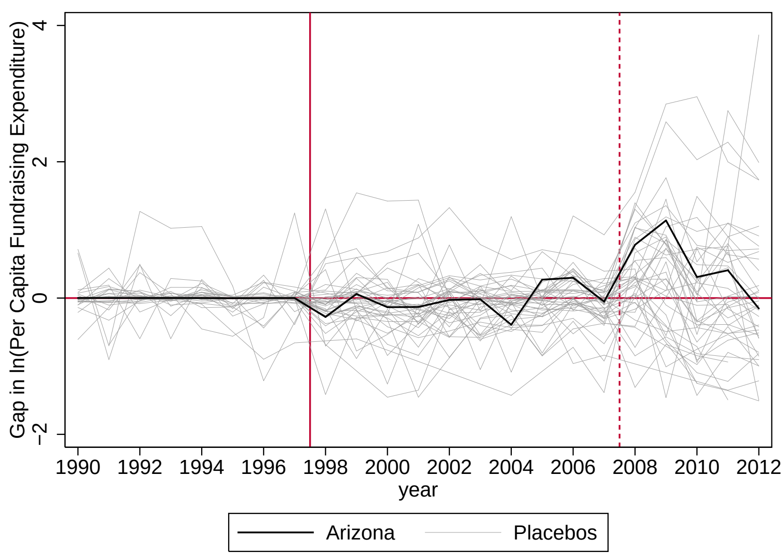
(b) Difference in the Log quantity of Community Foundations Per Capita relative to Synthetic Control: Iowa and Placebos

The Iowa time series in panel (a) represents the number of community foundations who filed IRS Form 990. Synthetic Iowa is derived from a donor pool of 39 untreated states. In panel (b), the black line displays the difference between Iowa and expected Iowa. Gray lines represent placebo tests. In both panels, the solid vertical line represents the introduction of Endow Iowa.

Figure 1.12: Synthetic Control Analysis: WPTC (Fundraising)



(a) Log of Per Capita Fundraising Expenditure by Six National Nonprofits, in Arizona and its Synthetic Control



(b) Difference in the Log Per Capita Fundraising Expenditure relative to Synthetic Control: Arizona and Placebos

The Arizona time series in panel (a) represents total per capita fundraising expenditures reported by six national nonprofits. Synthetic Arizona is derived from a donor pool of 43 untreated states. In panel (b) the black line displays the difference between Arizona and expected Arizona. Gray lines represent placebo tests. In both panels, the vertical lines represent the introduction of the WPTC (solid) and the change in how fundraising expenditure is reported on IRS Form 990 (dashed).

Chapter 2

Does AmeriCorps Crowd Out Private Giving?

2.1 Introduction

Since its creation in 1994, politicians have debated both the expansion and elimination of the AmeriCorps program. The federally-funded program offers stipends and scholarships to its members in exchange for their service at public or nonprofit organizations (Corporation for National and Community Service, 2009). Those pushing for expansion of the AmeriCorps program had a victory in 2009 with the passage of the Edward M. Kennedy Serve America Act (SAA). The law would have expanded funding for the Corporation for National and Community Service (CNCS) and increased the number of AmeriCorps members from 75,000 to 250,000 annually. This expansion, however, has not been fully funded. Instead, a 2013 report of the Budget Committee of the House of Representatives derided the agency for creating "the oxymoron–'paid volunteer'" while proposing the elimination of CNCS entirely (U.S. House Committee on the Budget, 2013). While funding for CNCS and AmeriCorps has held relatively steady, the President's 2018 budget again calls for elimination of

CNCS (Office of Management and Budget, 2017).

This paper examines the effect that AmeriCorps has on private charitable giving. This is done with two goals: to uncover hidden costs or benefits associated with a politically controversial federal program, and to provide, more generally, an estimate of the extent to which government funding for labor "crowds out" (reduces) or "crowds in" (increases) private giving. To do so, I investigate the relationship between AmeriCorps and contributions to nonprofits before and after two major changes in AmeriCorps policy.

There are many studies that estimate the extent to which government, monetary grants crowd out or crowd in private giving (Abrams and Schmitz, 1984; Kingma, 1989; Payne, 1998; Khanna and Sandler, 2000; Heutel, 2014; Duncan, 1999; Gruber and Hungerman, 2007). However, the empirical evidence has been mixed, likely because of the difficulty in disentangling the multiple forces at play. Grants may change the way that donors feel about the nonprofit's mission, or what they believe about its finances. Similarly, nonprofit managers may change fundraising or programmatic decisions based on whether or not they receive a grant. Public goods theory suggests that crowd out should be large, but imperfect information introduces the possibility that government grants can act as signals of the quality of nonprofits and increase giving at the firm level. Summarizing the empirical literature, Tinkelman (2010) finds that estimates of the level of crowd out form a bell curve centered just below zero; on average, crowd out exists but is low.

To date, few studies of crowd out have focused on labor, and those that have focused on labor estimated the relationship between government spending and volunteer labor (Day and Devlin, 1996; Menchik and Weisbrod, 1987; Duncan, 1999). This paper is novel because it estimates the relationship between government labor and private giving. Understanding the potentially disparate effects of grants of labor and capital on the decisions of nonprofits will give greater insight into the technology and constraints underlying public good production. Given the mixed evidence on crowd-out to date, studies like this one, that focus on specific types of grants, help to disentangle the circumstances in which crowd-out or crowd-in effects are dominant.

Accounting for the possibility that the number of and placement of AmeriCorps is jointly determined with donations is the major empirical challenge of this study. I address this issue by focusing on high frequency within-firm variation and an instrumental variables (IV) approach that exploits changes in AmeriCorps policy. The analysis is built upon administrative data on the number and placement of AmeriCorps members and financial information from the tax returns of 501(c)3 nonprofit organizations. The combination of these two data sources leads to a panel of charities for which donations, grants, and the number of AmeriCorps are known. Two pieces of legislation passed in 2009 altered the level of funding for AmeriCorps. The American Recovery and Reinvestment Act (ARRA), provided a one-year increase in AmeriCorps VISTA membership of 40 percent while the SAA increased funding levels for AmeriCorps State and National Direct. The instruments interact the treatment period of each policy with the level of AmeriCorps at a nonprofit sponsor prior to the legislative changes (2009).

I find evidence that AmeriCorps State and National crowd out contributions to their nonprofit partners. While the SAA increased the total number of AmeriCorps State and National, the policy also increased churn. First stage estimates imply that nonprofits with the most AmeriCorps in 2009 saw the largest reductions in AmeriCorps after implementation of the SAA. Among those nonprofits affected by this change, I find that a ten percent increase (decrease) in the number of AmeriCorps State and National is associated with a one percent decrease (increase) in contributions.

In contrast, estimates of the impact of the VISTA program are not suggestive of crowd out. I find that the additional VISTA members funded by the ARRA were awarded, disproportionately, to sponsors with the largest number of VISTA members at the time. Among nonprofits affected by the ARRA, estimates rule out—with 95 percent confidence—the possibility that a doubling of the number of VISTA would reduce contributions by more than three percent.

There are a number of potential policy benefits of the AmeriCorps program that this paper will not address. For example, if voters do not believe that nonprofits efficiently manage their resources, in-kind transfers may be more popular than grants of money. While the AmeriCorps program does provide funds to nonprofits, the regulations in how those funds are spent makes the program, in effect, an in-kind transfer of labor.[1] Additionally, the AmeriCorps program may encourage volunteerism and civic engagement in its membership that could be seen as beneficial beyond the term of service. This analysis ignores these long run effects and focuses solely on the impact the program has on the finances of its nonprofit partners.

In the next section, I frame this study in the context of the growing literature on crowd in and crowd out. As this is the first academic study to use the AmeriCorps data, I discuss it along with the NCCS dataset and the methodology employed to map the two together in detail. This is followed by a discussion of the estimation strategies. Results are reported in section 2.5 and are followed with a conclusion that discusses the implications of this research.

2.2 Background and Literature

The question of crowd out is interesting in regards to both efficient government spending and the motivations behind philanthropic giving. The classical economic theory predicts that, if donors are motived by consumption of public goods, grants crowd out

[1] AmeriCorps funding described in greater detail in the appendix. The VISTA and State and National Direct programs involve monetary grants that must be spent on the AmeriCorps program. The National Civilian Community Corps (NCCC) program, which I do not analyze in this paper, is directly government funded and would qualify as a system of in-kind labor grants.

private giving by as much as dollar-for-dollar (Bergstrom et al., 1986). More nuanced models allow for individuals to enjoy the act of giving itself. In these "warm glow" models, the level of crowd out shrinks to zero as donors' motivations shift from public good demand to the warm glow of giving (Andreoni, 1988, 1989).

Explanations for crowd in tend to assume that grants signal something positive about the nonprofit organizations that receive them.[2] Vesterlund (2003) and Potters et al. (2005) provide a theoretical model and experimental evidence to support a signalling hypothesis. Huetel (2014) finds that government grants have a greater crowd-in effect on younger organizations and explains that this trend is consistent with the imperfect information assumptions of the signalling model.

The existence of a crowd out effect does not necessarily imply that the government grant was the direct cause of the decline in private donations. There is growing evidence that nonprofit organizations reduce fundraising activities after receiving grants (Andreoni and Payne, 2003; Andreoni and Rao, 2011). Even if nonprofits maximize donations, Name-Correa and Yildirim (2013) show that, at least when fundraising is costly, government grants should reduce fundraising expenditure.

Empirical evidence on crowd out and crowd in is mixed. The majority of the research finds evidence of partial crowd out (Abrams and Schmitz, 1984; Kingma, 1989; Payne, 1998; Duncan, 1999; Gruber and Hungerman, 2007). However, Breman (2006) found crowd out to be near zero, and a number of more recent studies have found evidence of crowd in (Okten and Weisbrod, 2000; Khanna and Sandler, 2000; Andreoni et al., 2014; Heutel, 2014). Tinkelman (2010) summarizes 134 results from 46 different studies and describes the distribution of results as "a bell curve centered on crowd-out slightly above zero." (p.24).

While the majority of the crowd out literature has focused on capital rather than labor, there has been some examination into whether government spending crowds

[2] Duncan's (2004) novel "impact philanthropist" model also allows for government spending to crowd in private donations.

in or crowds out volunteer labor. Multiple studies find a crowd in effect, although when subdivided by sector, results are inconsistent (Menchik and Weisbrod, 1987; Day and Devlin, 1996). Duncan (1999), on the other hand, finds that government spending significantly crowds out gifts of both time and money. This paper examines the opposite relationship, whether the public provision of labor crowds in or crowds out private gifts of money.

Why would publicly provided labor be different than publicly provided money? From a donor's perspective, AmeriCorps may provide a signal that is different than the one provided by monetary grants. Yet, whether AmeriCorps provide a stronger or weaker signal of nonprofit quality is unclear. Even if the signal itself is the same, donors may be more aware that a nonprofit has AmeriCorps than they would be that it received a different type of grant. Additionally, there is evidence that donor responses are affected by how the public contribution is framed (Eckel et al., 2005). If donors perceive the funding mechanism behind AmeriCorps differently than the funding mechanisms behind other forms of grants, the crowding out effect is likely to be different.

Publicly provided labor may also affect the nonprofits in a unique manner. Revenue in the form of grants can be used by nonprofits to invest in physical capital, to spend directly on services, or to hire additional staff. Additional labor in the form of AmeriCorps, however, is less fungible and cannot be directly converted to capital or to defraying service costs. As such, AmeriCorps may impact nonprofits differently depending on their capital needs and constraints. Secondly, relatively little is known about the substitutability between employed and volunteered labor within a nonprofit (Handy et al., 2008; Simmons and Emanuele, 2010).[3] Where AmeriCorps fit in to the spectrum between employed and volunteer labor is even more unclear. Finally,

[3] Handy et al. (2008) find that only about 12% of tasks at nonprofits can be interchangeably performed by either volunteers or staff. Simmons and Emmanuelle (2010) find that volunteers are substitutes for low-wage workers. These findings are generally consistent and a first step, but tell us very little about skilled workers and volunteers.

fundraising professionals may respond differently to AmeriCorps than they would to additional revenue. If AmeriCorps do the same work as nonprofit professionals, their presence may reduce budgetary goals and fundraising targets. AmeriCorps might be employed in a way that allows a nonprofit to spend more resources on fundraising, or the grants management process may be so onerous that it detracts from organizational fundraising capability.

Few studies examine the relationship between national service programs and their partner organizations.[4] Existing research on AmeriCorps focuses on the impact of the program on its members. There is both non-causal and quasi-experimental evidence that AmeriCorps members become more involved in their communities after service (Simon and Wang, 2002; Simon, 2002; Frumkin et al., 2009). While it is not known whether AmeriCorps service has a causal relationship with future civic engagement, it is worth keeping in mind that there could be long term effects on private philanthropy through AmeriCorps alumni that I do not estimate.

This paper builds on the studies above and makes a contribution to the larger literature about the interaction between the government and the nonprofit sector. By examining the relationship between AmeriCorps and private philanthropic donations, it is novel in two ways. First it examines the effects of grant-funded labor on private philanthropy. Second, it studies the impact of major national service programs on the finances of nonprofit organizations.

[4] Brudney and Gazley (2002) find that the Small Business Administration's all volunteer Service Corps of Retired Executives led to increased services and no cutbacks in personnel. However, there is no reason to assume that evidence from the public sector would be a good predictor of the actions of private nonprofits.

2.3 Data

2.3.1 AmeriCorps Data

There are three primary AmeriCorps programs: AmeriCorps State and National Direct, Volunteers in Service to America (VISTA), and the National Civilian Community Corps (NCCC). Additionally, AmeriCorps are granted through AmeriCorps Tribes and Territories and, until 2012, Fixed Amount Grants.[5] However, these programs are effectively the same as State and National Direct programs and are treated as such throughout this analysis. The NCCC program places its members at regional campuses and then assigns teams to short-term service projects throughout the region. For each of the other programs, AmeriCorps members work exclusively with a sponsor organization—many of which are 501(c)3 nonprofits.

Appendix A provides additional information on the history of AmeriCorps and the process by which the program is funded, sponsor organizations—nonprofits or local government agencies—are chosen, and members join. It is important to note that AmeriCorps members are either paid or given a stipend for their service, that they are eligible for either a scholarship or a cash bonus after completing their term of service, and that funding mechanisms differ from program to program. While the scholarship (formally the Segal Education Award) is federally funded, funding for salaries or stipends may come from either federal or state service commissions or from the sponsor organization. The VISTA program provides more federal funding on a per-member basis—about $18,000 per member compared to $9,000 per State and National Direct member.

The data on AmeriCorps appear in publicly available "Full Reports" for each state and for each program year from 2004/05 to 2012/13. The reports include information

[5] Fixed Amount Grants are listed as "Education Awards Program" from 2005 to 2010 and "AmeriCorps Fixed Amount Grant" for 2011 and 2012.

on AmeriCorps, Senior Corps, and the Learn and Serve America program.[6] Program years run from summer to summer as many AmeriCorps programs are aligned with academic calendars. Since the relationship of interest is the impact of AmeriCorps on donations (rather than the other way around) the number of AmeriCorps in 2006/07 are matched with financial data for fiscal year 2007.

For each AmeriCorps grant, the Full Report lists the primary city, program or project name, sponsor organization, program type, and number of participants (AmeriCorps members). AmeriCorps NCCC members are not directly embedded with sponsor organizations and are, therefore, excluded from this analysis. Of the more than 50,000 AmeriCorps each year, fewer than 1,500 are members of NCCC. While the Full Reports are comprehensive in the enumeration of AmeriCorps members placed each program year, they have limited information on the sponsor organizations. AmeriCorps programs are listed in the reports by city or town and sponsor organizations are described only by name.

Additional data on the level of CNCS funding at the sponsor level, sponsor employee identification numbers (EINs), and sponsor addresses, was provided directly by CNCS. With the aid of research assistants funded through a grant from CNCS, this data was merged with the data from the full reports to create a panel of AmeriCorps sponsors. Organizations were matched by name and state, but discrepancies in naming conventions led to an imperfect match between the two data sources. In many cases, we were able to account for these inconsistencies, but it is possible that some errors remain and that any irregularities in the naming conventions used in the reports could have flowed through to the dataset.[7]

AmeriCorps sponsored by organizations that functioned in more than one loca-

[6] Full Reports can be downloaded from `www.nationalservice.gov/impact-our-nation/state-profiles` in PDF format.

[7] We identified 7,329 unique AmeriCorps sponsors from the state reports. The data provided by CNCS included 5,192 grantees, of which 3,676 could be matched to sponsors listed in the state reports. Additional research allowed us to add EINs to 159 nonprofits that appeared in the Full Reports but we could not match to the CNCS data.

tion were aggregated together and placed, where possible, in the city and state in which the nonprofit entity was headquarted and from which it would report its charitable receipts. Notably, this meant aggregating all of the AmeriCorps members in University of Notre Dame's ACE Leadership programs to Notre Dame, Indiana; all members serving with the Catholic Volunteer Network to Takoma Park, Maryland; and all members serving with the American Red Cross to Washington, DC.

The number of participants listed in the Full Report is the amount of AmeriCorps positions granted to the sponsor, not the number of positions filled.[8] Organizations do not receive funding for unfilled positions. The results presented in this paper can therefore be interpreted as estimates of intent-to-treat effects. From a policy perspective, the impact of the grant award is the relevant question and, therefore, the number of positions granted is the preferred explanatory variable. With regard to the direct effect of AmeriCorps on donations, these estimates should be seen as a lower bound.

Table 2.1 provides summary statistics for AmeriCorps grants and shows that grants vary greatly in size. The average State and National Direct grant provides 49 AmeriCorps members and the average VISTA grant provides 6 AmeriCorps members. The numbers, however, are highly skewed. Some very large grants were given to large, notable entities. The median number of AmeriCorps members by sponsor organization was only 20, less than one half the mean, and one fourth of all VISTA grants were for a single AmeriCorps member. Figure 2.1 graphs the number of AmeriCorps over the sample period. The impact of two legislative changes are visible. First, the American Recovery and Reinvestment Act of 2009 (known as ARRA or "the stimulus package") included increased funding for AmeriCorps that lead to a one-year increase in VISTA membership of more than 40 percent. Second, the Kennedy Serve America Act of 2009 (SAA) increased the number of State and National Direct by more than

[8] With the exception of prospective NCCC AmeriCorps, all other prospective AmeriCorps members apply to serve on a specific project with a specific organization.

10 percent beginning in 2011.

2.3.2 The Nonprofit Dataset

Data on nonprofit revenues comes from the Urban Institute's National Center for Charitable Statistics (NCCS) which compiles and harmonizes IRS data from nonprofit tax returns. I use the NCCS Core files for public charities (Core PC) from 2004 to 2013 to construct a panel of charities.[9] The data includes information on revenue and expenses including the amount received in private donations, government grants, membership dues, and program service revenue.[10] Additionally, NCCS has added classifications from the National Taxonomy of Exempt Entities (NTEE) for each organization which divides charities into both narrow and broad categories based on the type of service that they provide. For example,broad categories include "Education", "Arts, Culture, and Humanities," and "Mental Health and Crisis Intervention" while narrow categories include "Two-Year Colleges", "Symphony Orchestras", and "Substance Abuse Treatment".

Two steps are taken to clean the NCCS data before merging in the AmeriCorps Data. First, for any organization that filed multiple tax returns under the EIN, I keep the highest values of donations, revenue, and government grant receipts among the group.[11] Second, I add together the financial variables of organizations that share a name and address but have multiple EINs. This consolidates the financial data of organizations such as Volunteers of America, which filed 990 forms for 134 unique EINs from its national headquarters in 2011. Finally, I aggregate organizations with the same name and different addresses to either the state or city level.[12]

[9] Each annual file contains data from the IRS Form 990 in the year in which they were filed. The sources for the Core PC data are the IRS's Return Transaction files, Statistics of Income sample files, and 990 forms on GuideStar (www.guidestar.org).

[10] Program service revenue refers to revenue brought in to a nonprofit in exchange for services it provides. For example, this classification includes fees for medical services at nonprofit hospitals.

[11] Duplications in the data appear to be the result of redundant or amended tax filings.

[12] This step is necessary in order to match the NCCS data with the AmeriCorps data, because

While the NCCS dataset does not include all charitable donations made in a given year, the subset of nonprofits included within it accounts for more than 60% of all donations in a given year.[13] The missing data stems from charities that did not file tax returns with the IRS. Since organizations with more than $25,000 in gross receipts are required to file by law, the missing data corresponds generally with smaller organizations who would also be less able to meet the requirements of an AmeriCorps sponsor.

I map the dataset of AmeriCorps to the NCCS dataset first by EIN and then by organization name and state of operation using STATA's "merge" and "reclink" commands. The reclink command creates a match value based on the similarities between a series of string variables, in this case the spellings of the city, the state, and the organization name. After reviewing the results, I keep matches rating greater 0.999835, adding 8 additional matches. When non-sponsors are used as a control group, both the inclusion of false matches—that is assuming that AmeriCorps were in place where they where not—and the exclusion of true matches—missing AmeriCorps where they actually are—would bias the estimates toward zero.

For the primary analysis, I construct a balanced panel of nonprofits operating from 2007 to 2013. I exclude firms with zero or negative levels of contributions and those for whom contribution, program, or total revenue data is missing. Because the analysis focuses on the impact of the AmeriCorps program, the sample is reduced to include only those narrow NTEE classifications in which at least one nonprofit within

---

the AmeriCorps data does not include addresses. Ideally, organizations would be aggregated to the level at which resources are shared, fundraising efforts are organized, and donors contribute. However, with 2,677,670 observations, the dataset was too large to choose between state and city aggregation on a case-by-case basis. Based on research into a subset of nonprofits, and discussions with individuals working in the nonprofit sector, I aggregate affiliate organizations to the state level with three exceptions: Boys and Girls Clubs, Habitat for Humanity, and Big Brothers Big Sisters. These organizations are part of national networks, but since they are managed locally, they are aggregated at the city level.

[13] Total donations in the NCCS dataset range from 147 to 196 billion dollars annually. Donations in the NCCS dataset totalled $191 billion in 2011; total giving that year was $298 billion (Nichols, 2012).

the classification sponsored AmeriCorps between 2007 and 2013.

The resultant sample is an unbalanced panel consisting of 63,849 nonprofits and 446,943 annual observations. Table 2.1 displays summary statistics for AmeriCorps Grants in the final sample. The mean size of an AmeriCorps grant is slightly larger (50.2 rather than 49.4) than in the original dataset. Table 2.1 displays the level of donor contributions and total revenue for nonprofits in the panel. Nonprofits that sponsor AmeriCorps have much higher contribution and revenue levels, on average, than non-sponsors. There are two reasons to expect that this would be the case. First, the process of applying to sponsor AmeriCorps probably favors larger organizations that can hire professional grant-writers and create AmeriCorps programs with more members for fewer government dollars. Secondly, there may be some advantage to having a recognizable name or larger organizations may be able to lobby CNCS and state service offices directly.[14]

2.4 Estimation Strategy

The aim of this paper is to uncover the relationship between AmeriCorps and private philanthropy. The variable of interest is the charity's total income from private donors. The NCCS dataset includes annual data on contributions, from which I deduct the value of funding received from CNCS. (Nonprofits generally receive CNCS funds in the fiscal year preceding the AmeriCorps term.) I estimate distinct effects for the State and National and VISTA programs for two reasons. First, because the roles of the members are different—State and National work in direct service while VISTA work to build organizational capacity—and second, because they were affected by unique policy changes. In order to differentiate these effects, the baseline estimate includes unique coefficients for the number of State and National AmeriCorps, SN and the number of AmeriCorps VISTA, V.

[14] Conversations with nonprofit professionals suggests that this is sometimes the case.

The baseline estimation equation is therefore defined:

$$Contributions_{ist} = \alpha_i + \lambda_{st} + \beta_1 SN_{ist} + \beta_2 V_{ist} + \epsilon_{ist} \tag{2.1}$$

Charity-level fixed effects, α_i, are included to control for unobservable, time-invariant characteristics of the charity.[15] Sector-by-year fixed effects (using broad NTEE categories), λ_{st}, are included to control for changes in philanthropic tastes that may be correlated with changes in the funding priorities of CNCS. Standard errors are clustered at the state level.

Under this specification the coefficients β_1 and β_2 estimate the marginal relationship between contributions and changes in the number of AmeriCorps State and National and VISTA. Given the highly skewed distributions of both contributions and the number of AmeriCorps per sponsor, both variables are transformed in the baseline estimate using the inverse hyperbolic sine (IHS) function, $\text{arsinh}x = \ln(x+\sqrt{x^2+1})$.[16]

The methodology described above estimates a causal relationship if, and only if, the variation in the number of AmeriCorps sponsored by a nonprofit is exogenous or the sector-by-year fixed effects fully account for any unobserved factors—such as public good demand—that influence both the number of AmeriCorps and the contribution level. However, it is possible that AmeriCorps grantors favor nonprofits that they expect to raise additional revenue in the next year or that there is a correlation between the ability of nonprofit professionals to "win" AmeriCorps grants and to raise money from donors. If this is the case, the estimates obtained from equation 2.1 will be biased upward.

[15] Charities are aggregated to the level at which they engage in fundraising activities and report to the IRS.

[16] The advantage of a linear specification is that it allows for a test of one-for-one crowd out. If having an additional AmeriCorps is worth $10,000 to a nonprofit, one-for-one crowd out would correspond to a coefficient of -10. However, given the lack of prior empirical evidence for one-for-one crowd out, and the possibility that signalling, fundraising, or other non-linear effects may be at play, the loss of this test is not a major concern. Robustness checks using the linear specification produced results that were not statistically significant.

In order to address endogeneity, I employ an IV approach that exploits exogenous changes in federal policy that increased the number of AmeriCorps service members. The ARRA included a one-time, \$200 million dollar increase in funding for AmeriCorps which produced a 40 percent increase in the number of VISTA positions. Later that year, the SAA increased funding levels for the State and National Direct programs, beginning in the 2010/11 grant year.[17] In both cases, the policies produced higher numbers of AmeriCorps per approved sponsor.

Under the assumption that the SAA and ARRA created exogenous shocks to the AmeriCorps program, dummy variables for the time period affected by the policy change, $POST_{SN}$ and $POST_V$ can be used to create exogenous instruments. I interact these treatment dummy variables with the IHS of the number of AmeriCorps in the year in which the SAA and ARRA were passed. This produces two instruments—$POST_{SN} * arcsinh(SN_{is2009})$ and $POST_V * arcsinh(V_{is2009})$—to coincide with the two endogenous explanatory variables. The instruments estimate the differential impact of the policy change based on the number of AmeriCorps awarded for the 2008/09 grant year. In each case, a positive coefficient would indicate that the policy benefited existing or larger sponsors and a negative coefficient would indicate that the policy benefited new or smaller sponsors.

The underlying assumption in the IV analysis is that the differential effects of the policies on nonprofits with different numbers of AmeriCorps in 2009 is uncorrelated with trends in contributions except through the causal impact of AmeriCorps. This assumption is violated if some sectors have more AmeriCorps members, on average, than others and trends in contributions differ by sector. The sector-by-year fixed effects in equation 2.1, λ_{st} control for this possibility. Under the preferred specification, the relaxed assumption is that the differential effects of the policies on nonprofits with different numbers of AmeriCorps are exogenous after controlling for correlation

[17] I include more information on the legislative history of AmeriCorps is included in the Appendix.

between the number of AmeriCorps and sector.

2.5 Results

The first stage results for the IV analysis appear in table 2.2. Panel A estimates the relationship between the instruments and the number of AmeriCorps State and National. These estimates suggest that nonprofits who sponsored more State and National at the time that the SAA was passed saw larger proportional reductions in the number of State and National that they sponsored after the SAA was implemented. The baseline estimate, which includes the control for sector-by-year effects, appears in column 2. It implies the following: if nonprofit A had twice as many State and National as nonprofit B in 2009, we would expect them to have only 75 percent more State and National in 2011, 2012, and 2013. Conversely, the ARRA, which went into effect immediately following its passage, led to greater proportional increases in the number of VISTA at nonprofits that already sponsored the greatest number of VISTA. These results appear in panel B and imply that if nonprofit A had twice as many VISTA as nonprofit B in 2009, we would expect them to have one and half times as many VISTA in 2010.

Table 2.3 displays both least squares and IV estimates of the relationship between AmeriCorps and contributions. The preferred specification appears in column 2. OLS estimates in panel A show a positive relationship between AmeriCorps, of both types, and contributions. IV estimates in panel B provide evidence that State and National crowd out contributions.

The OLS specifications in Panel A of table 2.3 estimate the expected change in contributions given an endogenous change in the number of AmeriCorps. For nonprofits with large numbers of AmeriCorps, a 10 percent change in the number of AmeriCorps is associated with a 0.3 percent increase in contributions. With regards to larger changes, the coefficients suggest that nonprofits see an increase in contributions of about 3 percent with their first AmeriCorps member and increases of about 2 percent when the number of AmeriCorps doubles.

The IV specifications in panel B of table 2.3 provide estimates of crowd out. They estimate the impact that an exogenous change in the number of AmeriCorps has on contribution levels. I find a statistically significant crowd out effect from State and National. For nonprofits with large numbers of State and National, a 10 percent change in the number of AmeriCorps is associated with a 1 percent decline in contributions. With regards to larger changes, the coefficients suggest that a,nonprofit would see a decline in contributions of about 9 percent with their first AmeriCorps State and National and a decline of about 7 percent when the number of AmeriCorps State and National members doubles.

While the coefficient on the impact of VISTA is not statistically significant in the baseline specification, it is possible to rule out large crowd out effects. The 95 percent confidence interval around the coefficient in the baseline estimate is (-0.031, 0.149). The low end of the interval is suggestive of crowd out of only 2.7 percent for the first VISTA or 0.3 percent for a ten percent increase in VISTA members.

The difference between the OLS and IV results provide information about the influence of unobservable factors that endogenously affect both the level of AmeriCorps at a nonprofit and the level of contributions the nonprofit receives. In the case of State and National, the positive OLS estimate and negative IV estimate suggest that there are unobservable factors that are positively correlated with both member levels and contributions. There are at least two possible explanations for this. The first is that the preferences of CNCS and the state agencies responsible with awarding AmeriCorps are closely aligned with the preferences of donors. The second is that there are unobservable changes in nonprofit management talent associated with both an increased (decreased) ability to obtain AmeriCorps and an increased (decreased) ability to solicit philanthropic funds. The OLS and IV estimates for the VISTA program, however, are similar. This suggests that unobservable factors are either inconsequential or cancel each other out.

Alternative specifications appear alongside the baseline estimates in tables 2.2 and 2.3. The specification in column 1 excludes the sector-by-year dummy variables and produces similar OLS and IV estimates for State and National along with a statistically significant crowd in effect for the VISTA program. The primary rationale for the inclusion of the sector-by-year effects was the possibility that trends in which sectors receive more or fewer AmeriCorps members may be correlated with trends in which sectors receive more or fewer contributions. If this relationship is positive, estimates would be biased upward and may explain the larger coefficient found here. Columns 3 and 4 add state-by-year dummies and replace the sector-by-year dummies in the baseline with sector-by-year-by-region dummies. These additional controls do not substantially alter the results from the preferred specification.

The original panel is limited to nonprofits with narrow NTEE codes in which at least one nonprofit sponsors AmeriCorps. Column 5 in tables 2.2 and 2.3 excludes from the sample hospitals and universities, which tend to have high levels of contributions but rely disproportionately on earned revenue. The coefficients estimated using this sample are similar to those in the baseline estimate. Column 6 further limits the sample to nonprofits in narrow NTEE codes in which at least 10 percent of the nonprofits sponsor AmeriCorps. The OLS estimates of the relationship between State and National Direct and contributions is even larger than in the baseline estimate while the relationship between VISTA and contributions disappears. With the smaller sample size, the IV estimates suffer from weak instruments in the first stage. Finally, tables 2.2 and 2.3 show that the baseline IV estimates are not substantially altered if estimated with limited information maximum likelihood rather than two-stage least squares.

2.6 Discussion

As the merits of the AmeriCorps program continue to be debated, it is worth examining the impact the program has on the charities that it seeks to help. This paper explores the relationship between AmeriCorps and contributions to AmeriCorps sponsors. The IV results presented in panel B of table 2.3 imply that AmeriCorps State and National crowd out contributions. First stage IV results show that while the SAA increased the number of AmeriCorps State and National overall, the nonprofits who sponsored the most State and National at the time that the law was passed saw large proportional reductions in the number of AmeriCorps that they sponsored after the SAA was implemented. Among those nonprofits affected by this policy change, a 10 percent increase in the number of State and National caused a 1 percent decline in contributions ($2,169 at the mean of the distribution), or a 10 percent decline in the number of State and National caused a 1 percent increase in contributions. Conversely, while it is not clear whether VISTA have any impact on contributions, it is possible to rule out large levels of crowd out.

Despite the unique structure of the AmeriCorps program, the impact on nonprofits is similar to other sources of government grant funding. Both the estimated level of crowd out for AmeriCorps State and National and the statistically insignificant results for VISTA fall toward the center of the distribution of the empirical estimates described in Tinkelman (2010). This suggests that either AmeriCorps can serve as replacements for nonprofit staff, making the restrictions on AmeriCorps funding more fungible, or that the fungibility of grants is not a major factor in determining the level of crowd out.

It is important to note that these findings are unable to differentiate between the response of donors and the response of the nonprofit itself. An exogenous increase in the number of AmeriCorps might make donors more likely to give, but this effect could be offset by changes in the organizational structure and declines in fundraising efforts. For example, research by Andreoni and Payne (2011) finds that more than 70 percent of crowd out can be explained by declines in fundraising activities while discussions with nonprofits suggest that management of AmeriCorps takes considerable time and effort. This may explain the crowd out effect of AmeriCorps State and National, while the role of VISTA in capacity building may off-set these declines.

A related possibility is that nonprofits set fundraising targets to meet specific institutional goals, and that AmeriCorps are substitutes for traditional employees. At the mean, the crowd out induced by the first AmeriCorps State and National is equivalent to \$21,165. This is more than twice the cost of the average AmeriCorps to CNCS but might be the typical cost to a nonprofit of hiring a traditional employee to do the same job as an AmeriCorps. Future research into the AmeriCorps program, or other programs of grant-funded labor, could help to disentangle these effects. Better data collection or more transparency surrounding the grant funding decision might allow for future studies to better separate treated and untreated nonprofits.

Table 2.1: Summary Statistics

	Observations	Mean	Median	Std. dev.	Maximum
AmeriCorps Grants listed in State Reports, 2004/05 to 2012/13					
State and National Direct	11,232	49.4	20.0	155.4	6,109
VISTA	7,847	6.1	3.0	9.4	220
Total	18,659	32.7	10.0	123.5	6,109
Grant Award	6,533	0.455	0.192	1.359	32.750
AmeriCorps Grants in 2007-2013 Balanced Panel					
State and National Direct	2,106	50.2	20.0	194.3	5,603
VISTA	1,736	6.3	3.0	11.2	224
Total	3,609	32.6	9.0	151.2	5,603
Grant Award	2,435	0.477	0.168	1.754	32.750
Revenue levels in 2007-2013 Balanced Panel					
Contributions	446,943	2.369	0.193	22.927	2,127
Sponsors	3,609	14.855	2.066	76.010	1,712
All Other Nonprofits	443,334	2.267	0.190	21.946	2,127
Total Revenue	446,943	12.891	0.069	130.133	20,013
Sponsors	3,609	43.426	0.078	260.538	4,932
All Other Nonprofits	443,334	12.643	0.069	128.501	20,013

Notes: AmeriCorps data is derived from publicly available "Full Reports" for each state and for each program year from 2004-2005 to 2012-2013 which were downloaded from `www.nationalservice.gov/impact-our-nation/state-profiles`. Program years run from summer to summer. AmeriCorps NCCC members are not directly embedded with sponsor organizations and are, therefore, excluded from this analysis. Grant funding data provided by CNCS, inflated to 2013 dollars. The 2007-2013 balanced panel combines grants to the same sponsor organization and includes only those sponsors whose financial data was also found in the National Center for Charitable Statistics database produced from IRS Form 990 returns. The balanced panel consists of 63,849 nonprofits over 7 years for a total of 446,943 observations. The AmeriCorps sponsor subset presented here consists of 3,609 observations in which a nonprofit has sponsored AmeriCorps State and National or VISTA members. All monetary variables are displayed in million 2013 dollars

Table 2.2: First Stage Instrumental Variables Results

	(1)	(2)	(3)	(4)	(5)	(6)	(LIML)
Panel A: First Stage Estimates of SN_{ist}							
$POST_{SN} * arcsinh(SN_{is2009})$	-0.192***	-0.193***	-0.193***	-0.193***	-0.189***	-0.137*	-0.193***
	(0.039)	(0.039)	(0.039)	(0.039)	(0.043)	(0.065)	(0.039)
$POST_{V} * arcsinh(V_{is2009})$	-0.057	-0.057	-0.056	-0.056	-0.056	-0.106	-0.057
	(0.030)	(0.030)	(0.030)	(0.030)	(0.033)	(0.083)	(0.030)
F Test of Excluded Instruments	12.6	12.73	12.74	12.74	10.41	2.67	12.73
Angrist-Pischke F Statistic	25.68	26.01	26.06	26.03	20.3	9.42	26.01
Panel B: First Stage Estimates of V_{ist}							
$POST_{SN} * arcsinh(SN_{is2009})$	-0.007	-0.007	-0.007	-0.007	-0.005	-0.068**	-0.007
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.026)	(0.008)
$POST_{V} * arcsinh(V_{is2009})$	0.325***	0.324***	0.324***	0.324***	0.342***	0.203**	0.324***
	(0.035)	(0.035)	(0.035)	(0.035)	(0.036)	(0.075)	(0.035)
F Test of Excluded Instruments	52.32	52.09	52.01	52.36	46.26	15.68	52.09
Angrist-Pischke F Statistic	103.4	102.7	102.51	103.22	93.25	18.4	102.7
Fixed Effects (α_i)	✓	✓	✓	✓	✓	✓	✓
Year Dummies	✓						
Sector x Year (λ_{st})		✓	✓		✓	✓	✓
State x Year			✓				
Sector x Region x Year				✓			
Exclude Hospitals and Universities					✓		
High Intensity Subset						✓	
Observations	446943	446943	446943	446943	428897	11361	446943

* p< 0.05 ** p< 0.01 *** p< 0.001

Notes: Table displays first stage instrumental variables estimates associated with panel B of table 2.3. SN_{ist} and V_{ist} are the inverse-hyperbolic sine of the number of AmeriCorps State and National and VISTA. Regressions (1-6) estimated with two-stage least squares; standard errors clustered on state. The baseline specification appears in column (2). Regression (5) excludes hospitals and universities and regression (6) limit the sample to nonprofits in NTEE categories in which at least 10 percent of nonprofits sponsor AmeriCorps. (LIML) is the baseline specification estimated with limited information maximum likelihood estimation.

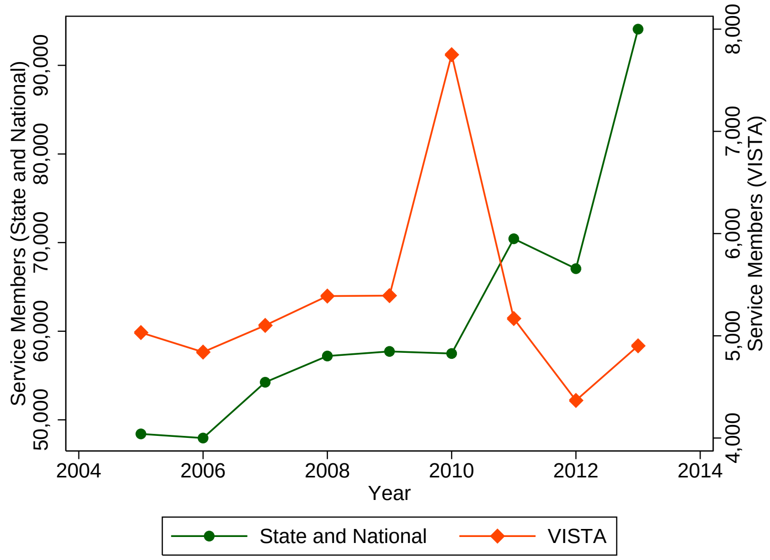
Table 2.3: Relationship between AmeriCorps Sponsorship and Contributions

	(1)	(2)	(3)	(4)	(5)	(6)	(LIML)
Panel A: Ordinary Least Squares Estimates							
State and National (β_1)	0.034**	0.034**	0.034**	0.035**	0.038***	0.047**	
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.016)	
VISTA (β_2)	0.029**	0.028*	0.027*	0.028*	0.032**	0.005	
	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.033)	
Panel B: Instrumental Variables Estimates							
State and National (β_1)	-0.098*	-0.109*	-0.108*	-0.110*	-0.089	-0.387	-0.109*
	(0.046)	(0.049)	(0.049)	(0.050)	(0.055)	(0.265)	(0.049)
VISTA (β_2)	0.092*	0.059	0.052	0.061	0.069	-0.106	0.059
	(0.046)	(0.046)	(0.045)	(0.047)	(0.048)	(0.171)	(0.046)
Kleibergen-Paap Wald F statistic	57.285	53.924	53.058	54.176	11.95	2.385	53.924
Kleibergen-Paap LM statistic	21.436	21.363	21.295	21.373	11.128	3.404	21.363
Fixed Effects (α_i)	✓	✓	✓	✓	✓	✓	✓
Year Dummies	✓						
Sector x Year (λ_{st})		✓	✓		✓	✓	✓
State x Year			✓				
Sector x Region x Year				✓			
Exclude Hospitals and Universities					✓		
High Intensity Subset						✓	
Observations	446943	446943	446943	446943	428897	11361	446943

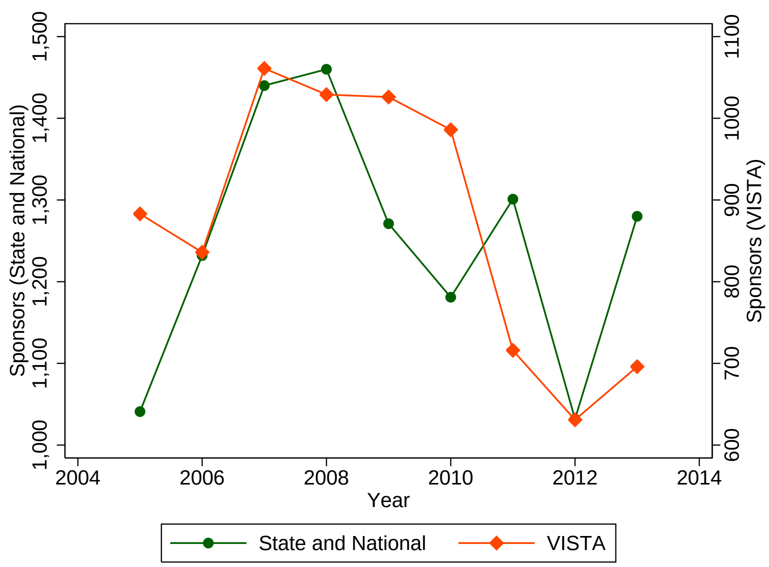
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Notes: Contributions are measured in real 2013 dollars. The regressions in the first four columns use the inverse-hyperbolic sine of contributions, the number of AmeriCorps State and National, and the number of AmeriCorps VISTA. Standard errors are clustered on state. In Panel B, regressions (1-6) are estimated with two-stage least squares; (LIML) is the baseline specification estimated with limited information maximum likelihood estimation. First stage estimates appear in table 2.2. The baseline specification appears in column (2). Regression (5) excludes hospitals and universities and regression (6) limit the sample to nonprofits in NTEE categories in which at least 10 percent of nonprofits sponsor AmeriCorps.

Figure 2.1: Annual AmeriCorps Membership and Sponsorship



(a) AmeriCorps Positions



(b) AmeriCorps Sponsors

Notes: AmeriCorps data is derived from publicly available "Full Reports" for each state and for each program year from 2004-2005 to 2012-2013 which were downloaded from www.nationalservice.gov/impact-our-nation/state-profiles. Program years frun from summer to summer. AmeriCorps NCCC members are not directly embedded with sponsor organizations and are, therefore, excluded from this analysis. The spike in the number of VISTA corresponds to increased funding from the American Recovery and Reinvestment Act. The increase in AmeriCorps State and National was funded by the Edward M. Kennedy Serve America Act.

Chapter 3

Cross-Country Inequality Analysis: How Sensitive are Results to the Choice of Data? (with Nora Lustig)

3.1 Introduction

There is a growing wealth of data that describes income inequality. There are at least fifteen international income inequality databases (herewith, databases) that contain a wide range of inequality indicators for many countries and many years (Ferreira et al., 2015). These databases are used for both academic and policy research (Piketty, 2014; Acemoglu et al., 2013; Ostry et al., 2014; Atkinson and Bourguignon, 2015). Yet, the inequality indicators contained within these databases differ in substantial ways. The differences in their methodologies and the effects of their methodological differences on estimates of inequality levels and trends are the subject of a series of papers included in a special issue of the Journal of Economic Inequality (Ferreira and Lustig, 2015). We focus here on nine databases that contain from 1,065 to 9,124 inequality indicators for many countries and many years.

This paper examines the extent to which differences between databases of inequality indicators defined by welfare concept, inequality measure, data source, country coverage, and time period affect inequality convergence-the finding that inequality has fallen in what had been highly unequal countries and risen in countries that had been more egalitarian (Benabou, 1996; Bleaney and Nishiyama, 2003; Ravallion, 2003). In particular, we analyze the sensitivity of the convergence result, both whether it exists and the rate at which it occurs, to choice of welfare concept (such as per capita consumption or equivalized household disposable income), inequality measure (Gini Coefficient, Theil Index, or Atkinson Index), the database used as source as well as the region over which convergence is estimated and the time period covered. This paper also estimates whether the rate of convergence has changed over time, and whether the results are more robust over some time periods than others.

Our data includes inequality indicators from nine sources. They are: *All the Ginis (ATG)*; United Nations Economic Commission for Latin America and the Caribbean (ECLAC)'s *CEPALSTAT*; the Organization for Economic Co-operation and Development (OECD)'s *Income Distribution Database (IDD)*; the Center for Distributive, Labor and Social Studies (CEDLAS) at National University of La Plata (UNLP)'s *Socio-Economic Database for Latin America and the Caribbean (SEDLAC)*; the *Standardized World Income Inequality Database (SWIID)*; *LIS* Cross-National Data Center in Luxembourg; UNU-WIDER's *World Income Inequality Database (WIID)* ; the World Bank's *World Development Indicators/POVCAL (WDI)*; and the *World Top Incomes Database (WTID).* Six of our sources (CEPALSTAT, OECD IDD, SEDLAC, WDI, and WTID) directly estimate inequality measures from microdata, two (ATG and WIID) collect indicators from other sources, and one (SWIID) uses a large number of sources to estimate annual series of probability distributions for the Gini coefficients using multiple-imputation methods.

The indicators included in our dataset estimate inequality across a variety of inequality measures and welfare concepts. While the Gini coefficient is the most frequent inequality measure, our analysis covers datasets that include the Atkinson Index, the Theil Index, the top one percent income share, top five percent income share, and the inverted Pareto-Lorenz coefficient. These measures are used to describe inequality in pre-tax income, disposable income, and consumption on both a per capita and equivalized basis. The databases that directly estimate inequality tend to have (broadly) consistent methodologies for defining their preferred welfare concepts. Those that do not-such as ATG and WIID-provide the researcher information with which he or she can decide how to construct panels with, hopefully, similar welfare concepts.

Our estimation framework follows Ravallion (2003) and estimates β convergence. Estimates of β convergence use regression procedures based on Ordinary Least Squares (OLS) to estimate the relationship between an initial level of inequality and the change in inequality over time. We test for inequality convergence by regressing observed changes in inequality on observed levels of inequality. We then test the sensitivity of the convergence results by holding the regression specification constant, while varying the panel of inequality indicators and performing a series of hypothesis tests to compare the estimates of the convergence rate produced by the different panels. Because measurement errors can bias the results towards convergence, we further follow Ravallion (2003) and use prior year's measurements of inequality as instrumental variables. It is important to note that we do not attempt to estimate a causal relationship between initial inequality and convergence or to explain the sources of inequality convergence. Moreover, we do not attempt to establish whether the convergence results are reflecting "true" convergence or just mere mean reversion. Rather, we focus specifically on how sensitive the convergence results are to the choice of inequality indicator.

With few exceptions, our estimates suggest that there is convergence in inequality levels. However, we find that the estimated rate of convergence varies significantly when we adjust our choice of data, with point estimates in the baseline ranging from converge at a rate of 11 percent to divergence at a rate of 2 percent. We find that the estimated rate of convergence is especially sensitive to the choice of which database is used as a source and which welfare concept is used from within a database. We also find that comparisons of the rates of inequality convergence across regions are frequently sensitive to the choice of welfare concept, time period, and database used in the analysis. Similarly, comparisons of the rate of inequality convergence across time are sensitive to the choice of database used as a source.

This paper continues as follows. Section 3.2 provides an overview of the literature that has explored the question of inequality convergence. We present our empirical model and the rates of convergence estimated by our iterative analysis in section 3.2. In section 3.4, we present our primary contribution: a series of hypothesis tests that examine the extent to which inequality convergence results are sensitive to the choice of welfare concept, inequality measure, database, country coverage (by region), and time period. Section 3.5 discusses our findings and concludes.

3.2 Inequality Convergence: A Brief Review of the Literature

Over the last 30 years, existing research suggests that inequality has fallen in what had been highly unequal countries and risen in countries that had been more egalitarian (Benabou, 1996; Ravallion, 2003). This "inequality convergence" appears to be part of a broader convergence of income distributions; as global inequality has declined, within-country levels of inequality have become more similar (Bourguignon, 2015).

As with any empirical study, however, the findings are subject to the strengths and weaknesses of the underlying data.

Global inequality convergence warrants further research for a number of reasons. First, the phenomenon is interesting in and of itself. It provides information on the distribution of income over time and explains the current trajectory of income inequality. By revisiting inequality convergence, we stand to gain a better understanding of whether the trends have changed. As levels of inequality have become more similar, has the rate of convergence slowed? Second, as explained by Benabou (1996) the neoclassical growth model predicts convergence in income distribution *among countries with similar fundamentals.* Therefore, analysis of convergence in income inequality, along with a parallel line of literature examining convergence in income *levels*, produces evidence that either supports or contradicts existing theory. Third, the information is useful to policy makers. Trends in inequality affect decisions about economic investment, redistribution, and trade policy.

Given its importance, there have been several examinations of inequality convergence. Benabou (1996) provides the first examination of inequality convergence using (then preliminary) data from Deininger and Squire (1996) and LIS. He finds evidence of inequality convergence in the 1970s, but little evidence of convergence in the 1980s, and no evidence of convergence over the entire two-decade period. Bleaney and Nishiyama (2003) provide further evidence of convergence using Gini coefficients reported in WIID 1.0. Bleaney and Nishiyama iterate their estimates using a"reliable data only" sample and a larger sample with"less reliable data" as well. Ravallion (2003) also finds evidence of convergence using data collected for an earlier project for the World Bank (Chen and Ravallion, 2001). He augments these results with a robustness check obtained using data from Li et al. (1998) based on Denninger and Squire (1996). Both Bleaney and Nishiyama (2003) and Ravallion (2003) find that while convergence itself is robust to the choice of database, the rate of convergence is highly sensitive.

Data reliability has been a central concern throughout the literature on inequality convergence. When Benabou (1996) first posed the question of inequality convergence he included the following caveat: "The paucity and sometimes poor quality of international data on income distribution remain binding constraints here, as in all empirical work on these issues," (p. 18). Similarly, Ravallion (2003) noted, "The shortage of comparable survey observations over time for many countries raises doubts about how well the trends have been estimated," (p. 355).

In an effort to limit data heterogeneity, a number of studies have focused on within-region or within-country inequality convergence across states or municipalities. Ezcurra and Pascual (2005) use data from the European Community Household Panel and find convergence within Europe. Gomes (2007) examines convergence inequality between Brazilian municipalities using data from the Joao Pinheiro Foundation's 2003 Brazilian Human Development Report. He finds evidence of general convergence and convergence within the southern region toward a lower inequality level while the rest of the regions converge to a higher inequality level. Panizza (2001), Ezcurra and Pascual (2009), Lin and Huang (2012) and Ho (2015) all focus on convergence between U.S. states. Ezcurra and Pascual (2009) use data from Patridge et al. (1996; 1998) based on Census Bureau estimates. Lin and Huang (2012) and Ho (2015) use data from Frank (2009) based on the Internal Revenue Service (IRS) Statistics of Income. (2001), Ezcurra and Pascual (2009), and Lin and Huang (2012) all find evidence in support of convergence in income inequality among U.S. states. Lin and Huang (2012) iterate their analysis with top 1% share, Atkinson Index, Gini Coefficient, relative mean deviation, and Theil Index. Their analysis shows that the general finding of inequality convergence among U.S. states is robust across inequality measures. However, their research does not examine differences in the rates of convergence between inequality measures. Ho (2015) is unique in that the results do not support the conclusion that income inequality levels among states are converging.

The reliability and heterogeneity in inequality data that has been a significant concern through out the convergence literature is the subject of our research. We test the robustness of the convergence result across sources, measures, and welfare concepts. Differences in our estimates are analyzed in light of differences in the underlying methodologies.

3.3 Estimating Convergence

We iteratively estimate inequality convergence by regressing observed changes in inequality on observed levels of inequality. We use a common specification, but vary the panel of inequality indicators to encompass alternative databases, welfare concepts, inequality measures, time periods, and country coverage. Our estimation framework follows Ravallion (2003) and estimates β convergence. Estimates of β convergence use regression procedures based on Ordinary Least Squares (OLS) to estimate the relationship between an initial level of inequality and the change in inequality over time.

There are a number of alternative tests for convergence. Ezcurra and Pascual (2005; 2009) examine the dynamics of the entire income distributions of European countries and US states. However, the Ezcurra and Pascual methodology can be applied only to highly homogenous data. Lin and Huang (2012) and Ho (2015) each present novel methods to tests for convergence based on panel unit root tests. Both studies, however, rely on a long panel of state-level US inequality data from Frank (2009) to estimate convergence in state-level inequality within the US. While methods based upon unit root tests could theoretically be used to estimate cross-national convergence, similarly appropriate data is not available. Unit root tests require datasets that are longer and more complete (meaning fewer years with missing data) than are generally found in the cross-national databases. Only SWIID provides such data at the national level, however, given that SWIID imputes every data point and employs a smoothing algorithm, we are not comfortable using it as a basis for this type of trend analysis.

Ezcurra and Pascual (2005) and Lin and Huang (2012), citing Quah (1993), criticize the econometric validity of the methodology employed by Benabou (1996), Ravallion (2003), and Gomes (2007) , that tests the prediction of convergence from neoclassical growth using OLS-based estimates of β convergence. The intuition, from Quah (1993), is that beta tests cannot distinguish convergence from mean reversion. As an example, suppose each country has some "baseline" level of inequality that may be increasing or decreasing along a long-term trend line. Further suppose there exists a number of one-time events that increase or decrease inequality initially, with inequality gradually returning to the long-term trend line. Our methodology does not distinguish between convergence in baseline inequality trends and reversion back to a pre-shock trend line. However, for our purposes, this issue is not fundamental in the following sense. Our study should be viewed as an exercise designed to assess the sensitivity of convergence estimates to the choice of welfare concept, inequality metric, time period, and database.

We address the possibility that measurement error could bias estimates of convergence through the use of IV. Suppose an inequality indicator, θ, at time $t = 0$ is measured with error ϵ_0. A change in inequality from $t = 0$ to $t = 1$ is then $\theta_1 + \epsilon_1 - \theta_0 - \epsilon_0$. OLS-based estimates of β convergence, such as those in this paper, compare the change in inequality, $\theta_1 + \epsilon_1 - \theta_0 - \epsilon_0$, to the initial level of inequality, $\theta_0 + \epsilon_0$. If ϵ_0 is positive (negative), the independent variable is biased upward (downward) while the dependent variable is biased downward (upward). This would, in turn, bias β downward leading to an overestimate (that is a more sharply negative estimate) of convergence. Following Ravallion (2003), we address this concern by using prior year's inequality indicators as instruments. In this example, the independent variable is now $\hat{\theta}_0(\theta_{-1} + \epsilon_{-1})$. As long as the errors are not serially correlated, the bias is removed.

While our methodology is not ideal for determining the causal rate of convergence, it is ideal for the comparison of panels of inequality indicators for the following reasons. First, since the focus of our work is to compare results across welfare concept, inequality measure, time, and databases, we want to use a methodology that resembles the most similar work: the tests for global inequality convergence by Benabou (1996), Bleaney and Nishiyama (2003), and Ravallion (2003). All three studies focus on convergence. Second, most of the cross-national databases provide either short (as in the case of OECD IDD) or incomplete (as in the case of WDI) time series. As such, these databases are a better fit for a methodology based on cross-sectional inference than the time series approaches that have been used on long panels (such as those used by Lin and Huang (2012) and Ho (2015)). The Ravallion methodology allows us to include countries with as few as three estimates of inequality. Third, convergence, by definition, estimates a convergence "rate". The Ravallion methodology, therefore, allows us to examine not only whether the convergence result is robust across alternative specifications, but whether the rate of convergence is sensitive specifically to alternate sources, welfare concepts, and inequality measures.

We begin by defining Θ to be a panel of inequality indicators. Within a panel of indicators, let years be indexed $t = (0, 1, \ldots, D)$ and let countries be indexed $i = (1, \ldots, N)$ such that the inequality indicator for a given country in a given year is denoted as θ_{it}. Following Ravallion (2003) we assume that there is some true level of inequality, θ^*_{it}, that is measured with error such that $\theta_{it} = \theta^*_{it} + \epsilon_{it}$.

Additionally, assume that each country has an underlying inequality trend $\tau_i(\theta^*_{i0})$ defined by the equation:

$$\theta^*_{it} - \theta^*_{i1} = \tau_i(t-1) + v_{it} \quad (i = 1, \ldots, N; t = 2, \ldots, D) \tag{3.1}$$

where v_{it} is a zero-mean, country and time specific error term that denotes true, short-term deviations from the time trend. Further assume that measurement error has a mean of zero and is serially independent such that $\theta_{it} = \theta^*_{it} + \epsilon_{it}$. Finally, assume that there exists some linear approximation of the inequality trend such that

$$\tau_i \approx \alpha + \beta\theta^*_{i1} + \mu_i$$

Then, we can rewrite equation 3.1 as:

$$\theta_{it} - \theta_{i1} = \alpha(t-1) + \beta\theta_{i1}(t-1) + e_{it}$$

$$(i = 1, \ldots, N; t = 2, \ldots, D) \tag{3.2}$$

where e_{it} is a composite (heteroskedastic) error term defined:

$$e_{it} = v_{it} + \epsilon_{it} - \epsilon_{i1} + (t-1)(\mu_i - \beta\epsilon_{i1})$$

Under the assumption that measurement errors are serially independent, we can use the inequality indicator from an earlier year, $\theta_{is}(s < 1)$ as an instrument for θ_{i1} and estimate equation 3.2 using an IV approach. Suppose, for example, we have estimates of the Gini coefficient for the years 1992, 1993, 1996, and 1997. We would use the Gini in 1992 as the instrument, denoting 1992 as $(t = 0)$ for the Gini in 1993 $(t = 1)$, and run our regression using observations of changes in inequality through 1995 $(t = 4)$ and 1997 $(t = 5)$. Similarly, if the Gini coefficient for 1992 is unavailable, the index for 1991 $(t = -1)$ could serve as an instrument. To maximize the strength of our instruments, we use the closest, available, prior estimate of inequality.

Using this methodology, we estimate β convergence iteratively, using each combination of welfare concept and inequality measure available in the nine databases listed in the introduction. Table 1 displays estimates for the period 1988 to 2012. In tables 2 and 3, the time frame is split in half, estimating convergence from 1988 to 2000 and from 2000 to 2012, respectively. In table 4, we use the same time frame as Ravallion (2003): 1983 to 1999. All estimates for the period 1988 to 2012 are negative (although not all are statistically significant) and therefore suggestive of convergence. For the time periods 1988 to 2000 and 1983 to 1999, we estimate *divergence* using OECD IDD's and LIS's measures of equivalized disposable income (although the LIS estimate is only statistically significant in the latter period). With this exception, however, our results generally confirm the existing findings in the literature.

These estimates provide a first pass at answering the question, "is there inequality convergence?" Generally, we find that there is. In the next section, we go further, and ask whether the rate of inequality convergence is sensitive to the choice of data. In our sensitivity analysis, we use the same databases, welfare concepts, and inequality measures. As such these estimates also provide a baseline for the pair-wise comparisons in section 4. We alter the data in only one way as we move from the tests of convergence presented here to the pairwise hypothesis tests that follow: the pair-wise tests use only data from countries that appear in both panels (i.e., both OECD and LIS disposable income or both WIID income and WIID consumption).

3.4 Hypothesis Testing

In order to understand how specific methodological and data choices affect estimates of convergence, we begin by defining inequality indicators along three dimensions. In addition to the typical distinctions of how welfare is defined (i.e., consumption or income and per capita or equivalized) and by which inequality measure (Gini, Theil, and so on) is used, given the main focus of our paper we also define each inequality indicator by the database from which it is drawn. Additionally, of course, each indicator is defined by a country and year. A panel of indicators, Θ, can similarly be described along the dimensions of the welfare concept, inequality measure, and source along with the set of countries and years included within it. Thus, let Θ_{wmjIT} be the panel of indicators using welfare concept w and inequality measure m, from database j, covering the set of countries I and time period T. Moreover, let $\hat{\beta}_{wmjIT}$ be the estimate of β convergence produced from panel Θ_{wmjIT}.

Using this notation, we set about testing a set of five "straw man" hypotheses associated with variation in welfare concepts, w, inequality measures, m, databases, j, sets of countries, I, and time periods, T, while holding everything else equal. In each case, we make pair-wise comparisons between panels using F-tests with the null hypothesis that the estimated rate of convergence, $\hat{\beta}_{wmjIT}$, would be the same using either panel. We limit our analysis to pair-wise comparisons in which each panel includes at least ten countries and at least 30 observations. We call these set of five "straw man" hypotheses A, B, C, D, and E; they are described in detail below. We then analyze the results of these five sets of hypothesis tests with particular regard to how different databases can yield different conclusions even in the cases in which the welfare concept, the inequality measures, the countries, and the time period are kept the same.

3.4.1 Hypotheses Set A: changing the welfare concept

We begin by estimating convergence with different welfare concepts that appear in the same database, while holding the inequality measure, country coverage, and time period constant. We then test the null that the estimated rates of convergence are identical. Formally our hypothesis is as follows:

Hypothesis A *Given any two panels of inequality indicators* Θ_{vmjIT} *and* Θ_{wmjIT}*, with identical inequality measures m, sources j, sets of countries I, and time periods*

T, the estimated rate of convergence is constant across welfare concepts, v and w; $\hat{\beta}_{vmjIT} = \hat{\beta}_{wmjIT}$.

We are able to test this hypothesis using six pairwise, within-databases comparisons. As an example, we compare estimates using SEDLAC's Theil Index estimates of equivalized income inequality to SECLAC's Theil Index estimates of per capita income inequality. Where there is sufficient data, these comparisons are made over the four time periods discussed in section 3.2—1988 to 2012, 1988 to 2000, 2000 to 2012, and 1983-1999—for a total of 20 hypothesis tests. Results appear in table 3.5 with the β coefficients plotted in figure 3.1 in declining order of the statistical significance of the F test between the two estimates.

In five out of the 20 cases, the null hypothesis of constant β coefficients across panels can be rejected at either a 5 percent or 1 percent significance level. We apply 12 tests of Hypothesis A using data from SEDLAC. In each case, we fail to reject the hypothesis that the estimated rate of convergence is identical whether the welfare concept employed is equivalized disposable income or per capita disposable income. In five of the other eight tests, however, we find that estimates of convergence are not equivalent across welfare concepts.

Using data from either SWIID or OECD IDD leads to the conclusion that, at least prior to 2000, the rate of convergence has been significantly higher when the welfare concept is based on pre-tax income than when it is based upon post-tax income. Using these two databases, we reject the null five times out of seven. As we noted in the prior section, using the time period from 1983 to 1999, data from OECD IDD suggests convergence in pre-tax and transfer income inequality and divergence in disposable income. The pairwise test in this case appears as the third test from the left in figure 3.1 and estimates a convergence at a rate of 6.6 percent using the pre-tax metric and divergence at a rate of 1.6 percent using the disposable income metric.

Examining our initial estimates created using WDI data, we find that estimates of income inequality convergence differ from estimates of consumption inequality convergence. For the period from 2000 to 2012 (table 3.3), the point estimate for the rate of income inequality convergence is -0.030 with a standard error of 0.004. In contrast, the point estimate for consumption inequality convergence is 0.009 with a standard error of 0.009. Yet, only a small subset of Eastern European Countries (Bulgaria, Estonia, Hungary, Latvia, Lithuania, Slovak Republic, and Slovenia) appear in both the income and consumption panels. As such, it remains unclear whether the different point estimates are driven by the choice of welfare concept or the variation in the countries included in the panels.

3.4.2 Hypotheses Set B: changing the inequality measure

Our second hypothesis estimates convergence with different inequality measures that appear within the same database and tests whether the estimated rates of convergence are equal. Formally our hypothesis is as follows:

Hypothesis B *Given any two panels of inequality indicators* Θ_{wljIT} *and* Θ_{wmjIT}*, with identical welfare concepts* w*, sources* j*, sets of countries* I*, and time periods* T*, but different inequality measures the estimated rate of convergence is constant across inequality measures,* l *and* m*;* $\hat{\beta}_{wljIT} = \hat{\beta}_{wmjIT}$

For example, under Hypothesis B, we test whether the estimated rate of β convergence is the same whether we use CEPAL estimates of per capita total current income inequality using the Gini coefficient or using the Theil index. Again, where data is available, we iterate our tests over four time periods and produce 33 unique tests of Hypothesis B.

Results for tests of Hypothesis B appear in table 3.6; β coefficients are plotted in figure 3.2. Even if alternative inequality measures constantly pointed to the same qualitative conclusions, we might find that estimates of the convergence rates are very sensitive to the choice of inequality measure. Yet, we are only able to reject the null of identical β coefficients in two of 33 cases. This result is consistent with the robustness checks presented in Lin and Huang (2012). The two instances of statistically significant difference occur in comparison between WTID estimates of the inverted Pareto Lorenz coefficient and the top five and top one percent income share. Here the estimates of convergence with the income share metrics are -0.0372 and -0.0256, respectively, while the estimates using the inverted Pareto Lorenz coefficient are 0.0007 and 0.0033.

3.4.3 Hypotheses Set C: changing the database used as a source

Our third hypothesis compares estimates of convergence using different databases. Here, we construct panels from pairs of databases such that the welfare concept, inequality measure, and country coverage are identical. We then compare, for example, rates of convergence estimated with LIS and OECD. We examine four different time periods and five different welfare measures to produce a total of 60 tests of Hypothesis C, stated formally as follows.

Hypothesis C *Given any two panels of inequality indicators Θ_{wmjIT} and Θ_{wmkIT}, with identical welfare concepts, w, inequality measures, m, sets of countries, I, and time periods, T, the estimated rate of convergence is constant across data sources, j and k;* $\hat{\beta}_{wmjIT} = \hat{\beta}_{wmkIT}$

Hypothesis C, which initially drove this research, posits that our estimates of convergence are unaffected by the choice of database, provided that the welfare concept and inequality measure are held constant. We run 60 tests of this hypothesis, all of which compare panels that employ the Gini coefficient, and reject the null 25 times.

F-statistics appear in table 3.7. The estimated rates of convergence, plotted in figure 3.3, vary from divergence at a rate of 7.8 percent (the WDI per capita consumption Gini coefficient from 2000 to 2012) to convergence at a rate of 9.3 percent (the WIID per capita expenditure Gini coeffient).

Our analysis shows that estimates produced using LIS and OECD are similar. We therefore cannot reject the hypothesis that the two databases yield the same estimated rate of convergence (provided we use an analogous indicator). LIS and OECD are similar in a number of ways. Both focus on economically advanced countries, calculate inequality from micro-data, and provide estimates of equivalized household disposable income. Their major difference is that LIS standardizes microdata from income surveys in-house prior to calculating inequality, while OECD IDD calculates inequality in conjunction with national statistical offices.

Comparison between SWIID and its source material provides interesting results. SWIID imputes complete time series using a variety of inputs including LIS, OECD IDD, SEDLAC, WDI, ATG, and WIID. The final SWIID time series is built to impute "LIS-comparable net income inequality" indicators (Solt, 2016). Additionally, SWIID employs a moving average formula to avoid "unrealistic" jumps in inequality from one year to the next-unless those jumps are documented by LIS. Yet, using three of the four time periods for our panels, we reject the hypothesis that estimates of convergence in equivalized disposable income are constant between SWIID and LIS. Focusing on the net income inequality indicators, we similarly reject the hypothesis that estimates of convergence in equivalized disposable income are constant when comparing SWIID with WIID over two of the four time periods and when comparing SWIID with OECD IDD over three of the four time periods. Conversely, we find that estimates produced using SWIID are generally similar to those produced using SEDLAC (with the 1983 to 1999 time period an exception). We also fail to reject the null in any of the comparisons between the SWIID and OECD IDD panels of pre tax and transfer income inequality.

When we test for the equality of estimates produced from panels that aggregate various welfare concepts together, we reject eight of 12 tests. The income and consumption inequality estimates in WDI are estimated by the World Bank using either grouped or micro data. ATG and WIID, however, aggregate Gini estimates from multiple sources. Panels constructed using these sources, therefore, include both aggregated welfare concepts, and aggregated methods of treating the microdata. As such, these rejections are consistent with the rejection of five of 12 tests of Hypothesis A which suggests variation in estimates created using alternative welfare concepts within the same IDD.

3.4.4 Hypotheses Set D: changing the region of analysis

Hypothesis D examines whether the estimated rate of convergence is constant across regions of the world. Here there are two questions of interest. First, is there variation in convergence across regions? We are interested in, for example, whether countries within Latin America and the Caribbean are converging at the same rate as countries in Sub-Saharan Africa. Second, do different databases produce different conclusions with regards to these sort of pair-wise comparisons? Specifically, the hypothesis can be stated as follows:

Hypothesis D *Given any two panels of inequality indicators* Θ_{wmjHT} *and* Θ_{wmjIT}*, with identical welfare concepts, w, inequality measures, m, sources, j, and time periods, T, the estimated rate of convergence is constant across regional country sets H and I;* $\hat{\beta}_{wmjHT} = \hat{\beta}_{wmjIT}$.

To test this hypothesis we make pair-wise comparisons between regions while holding the inequality measure (always the Gini coefficient), welfare concept, time period, and source database constant. We test Hypothesis D using the following sources and income concepts: WDI estimates of income inequality, WDI estimates of consumption inequality, WIID estimates of per capita disposable income inequality, SWIID estimates of market income inequality, and SWIID estimates of net market income inequality. By varying the time periods over which we estimate convergence, we produce a total of 191 pair-wise comparisons between the following regions: Advanced Economies, East Asia and the Pacific, Europe and Central Asia, Latin America and the Caribbean, the Middle East and North Africa, South Asia, and Sub-Saharan Africa. Given the focus of this paper, we then compare whether the results of the hypothesis tests vary whether we use WDI, WIID, or SWIID.

We reject the null of equal rates of convergence in 111 of 193 tests. This includes 15 tests using data from WDI, eight tests using data from WIID, and 168 tests using data from SWIID. We find that results vary even within databases when the income concept or time period varies. For example, using panels of WDI/POVCAL income inequality Gini coefficients, we compare rates of convergence between panels of Latin American countries with panels of advanced economies across four time periods. Strikingly, we reject the null of equal rates of convergence for the period 1988 to 2000, but fail to reject the null or the period 1983 to 1999. Similarly, for the period 1988 to 2012, the panel of WDI/POVCAL consumption inequality Gini coefficients leads to the rejection of the hypothesis that convergence rates are the same in Europe and Central Asia as in advanced economies, but fail to reject that hypothesis when using income inequality Gini coefficients. The 168 tests performed using SWIID include 103 in which we reject the null hypothesis. More often than not, estimated convergence rates differ by region. Yet this finding is highly sensitive to the specific panel employed to test it.

3.4.5 Hypotheses Set E: changing time periods

Finally, we test whether the estimated rate of convergence is constant across time. Here, we hold the source database, welfare concept, inequality measure, and group of countries constant while estimating the rate of convergence over two different time frames.

Hypothesis E *Given any two panels of inequality indicators* Θ_{wmjIS} *and* Θ_{wmjIT}*, with identical welfare concepts, w, inequality measures, m, databases, j, and sets of countries, I, the estimated rate of convergence is the same over time periods S and T;* $\hat{\beta}_{wmjIS} = \hat{\beta}_{wmjIT}$.

We make pair-wise comparisons between the time periods 1988 to 2012, 1988 to 2000, 2000 to 2012, and 1983-1999-six total comparisons-using 27 different sets of inequality indicators. In total we estimate 153 test of Hypothesis E. We then compare whether the hypothesis is accepted or rejected using each database.

We reject Hypothesis E, the proposition that estimated convergence rates remain constant as the time period under study varies, 57 out of 153 times. Here, there is frequently consistency across welfare concepts and inequality measures for a given comparison within a given database. For example, we consistently fail to reject the null that convergence rates from 1988 to 2000 were equal convergence rates from 2000 to 2012 when using SEDLAC and CEPAL while we reject null using either the Gini coefficient or the Atkinson Index contained in LIS. Surprisingly, however, the panels made up of SEDLAC or CEPAL data lead to rejection of the null hypothesis in five out of seven tests when we compare 1983 to 1999 with 2000 to 2012. As another example, WTID data allows for the rejection of the null that the rate of convergence of the top one percent of incomes was the same from 1988-2000 and from 2000-2012. However, we fail to reject the same null hypothesis using either the top five percent income share or the inverted Pareto Lorenz coefficient.

In summary, this analysis explores the sensitivity of the estimated rate of convergence by testing five hypotheses using a series of pair-wise F-tests. The main findings of this sensitivity analysis are summarized in table 3.8. The main takeaways are as follows. First, estimates appear to be more sensitive to the choice of welfare concept than to the choice of inequality measure. Second, different databases frequently produce different results, even when the countries, the welfare concept, the inequality measure, and the time period are held constant. Third, while there is a rather large amount of evidence that estimated rates of convergence differ by region and by time—Hypotheses D and EE are rejected 58 percent and 37 percent of the time—even this result is sensitive to the database that is used to perform the analysis.

3.5 Conclusions

The purpose of our paper is to assess the extent to which inequality analysis is sensitive to the choice of the data. We use iterative estimation of inequality convergence as a means to this end. We examine the sensitivity of our estimates of convergence to the choice of welfare concept, inequality measure, the database used as a source, as well as the region over which convergence is estimated and the time period covered.

We use data from nine databases with various welfare concepts, inequality measures, and country coverage. Overall, our estimates are generally supportive of the concept of inequality convergence (or mean reversion since-strictly speaking-we cannot distinguish between the two). Although not all of our estimates are statistically significant, we only rule out convergence using OECD IDD's measure of equivalized disposable income for the periods 1988 to 2000 and 1983 to 1999 and LIS's measure of equivalized disposable income for 1983 to 1999. These specific results are notable for two reasons. First, the fact that any of our estimates are statistically significant and positive provides a contrast with the existing literature. Second, for the period of 1988 to 2000 and 1983 to 1999 we find different results within the same database using alternate welfare concepts. That is, we find divergence (we reject a null of convergence) using OECD disposable income and we find convergence (we reject a null of no convergence) using OECD pre-tax and transfer income.

We examine this phenomenon more rigorously in our tests of Hypothesis A. We reject five of twenty tests in which we hold the time period, database, countries, and inequality measure constant, but vary the welfare concept. More specifically, where we compare welfare concepts that differ between pre-tax (or market) and post-tax (or disposable) income, we reject five of seven tests. On the other hand, using SEDLAC, we do not reject the null that estimated rates of convergence vary depending on whether the welfare concept is per capita or disposable income.

In contrast with our other results, cross-sectional estimates of inequality convergence appear to be relatively consistent across inequality measures. While the Gini coefficient is used most prominently, we find little evidence to suggest that estimates would change significantly if the analysis were built on panels of Atkinson or Theil indices nor do our findings suggest a significant differentiation between using the top one percent income share or the top five percent income share. We should caution, however, that these inequality measures provide different information about the distribution of income and that while we find little evidence to suggest that the choice of measure alters large cross-sectional analysis, this does not mean that the metrics are interchangeable.

Estimates are highly sensitive to the source of the data. Even when the welfare concept and inequality measure are the same, and the countries used in the analysis are identical, results occasionally differ. We would therefore recommend that, where possible, all cross-national inequality studies test the sensitivity of their findings across multiple sources. Alternatively, we ought to regard any finding based on a single database as preliminary.

Of these findings, the most interesting may be the sensitivity of the convergence result to the choice of welfare concept. Yet, it may come as no surprise to researchers who focus on the relationship between fiscal policy and inequality. The Commitment to Equity Project, for example, provides a series of country-by-country examinations of the relationship between fiscal redistributions and economic inequality (Lustig, 2016). One conclusion that can be drawn from this research, and others like it, is that fiscal redistribution is driven by country-specific institutional characteristics and politics (Lustig et al., 2014). As such we should expect that any analysis of trends in inequality would be sensitive to whether one uses market income, disposable income, or consumption as the primary welfare concept. Any analysis based on inequality trends should therefore be very precise about the welfare concept that is being used and we should remain cautious about use of inequality indicators based upon broadly defined or poorly understood welfare concepts.

Finally, our results display the importance of panel construction and its effect on results. While the depth and breadth of inequality data continues to grow, large gaps remain. Adjusting, even slightly, the time period being studied or the regions included can alter estimates by a statistically significant margin. As such, researchers ought to be very specific about the choices that they make in constructing panels of inequality indicators and very humble about the external validity of their results.

Table 3.1: Convergence Estimates 1988-2012

Source (j)	Welfare Concept (w)	Inequality Measure (m)	Slope Coefficient	Std error	Intercept Coefficient	Std error	N
CEPALSTAT	Per Capita Total Current Income	Gini Coefficient	-0.026	(0.004)	0.014	(0.002)	177
CEPALSTAT	Per Capita Total Current Income	Theil Index	-0.032	(0.005)	0.018	(0.003)	177
CEPALSTAT	Per Capita Total Current Income	Atkinson Index (1)	-0.024	(0.004)	0.010	(0.002)	177
LIS	Equivalized Disposable Household Income	Gini Coefficient	-0.010	(0.003)	0.004	(0.001)	86
LIS	Equivalized Disposable Household Income	Atkinson Index (1)	-0.005	(0.004)	0.001	(0.001)	86
OECD IDD	Equivalized Pre Tax and Transfer Income	Gini Coefficient	-0.037	(0.005)	0.018	(0.002)	201
OECD IDD	Equivalized Disposable Household Income	Gini Coefficient	-0.010	(0.002)	0.004	(0.001)	224
SEDLAC	Per Capita Disposable Household Income	Gini Coefficient	-0.038	(0.003)	0.019	(0.001)	220
SEDLAC	Per Capita Disposable Household Income	Theil Index	-0.045	(0.005)	0.023	(0.003)	220
SEDLAC	Per Capita Disposable Household Income	Atkinson Index (1)	-0.040	(0.003)	0.015	(0.001)	220
SEDLAC	Household Equivalized Income	Gini Coefficient	-0.037	(0.003)	0.018	(0.001)	220
SEDLAC	Household Equivalized Income	Theil Index	-0.044	(0.005)	0.021	(0.002)	220
SEDLAC	Household Equivalized Income	Atkinson Index (1)	-0.040	(0.003)	0.014	(0.001)	220
Top Incomes	Income	Top 1% Share	-0.009	(0.003)	0.205	(0.022)	340
Top Incomes	Income	Top 5% Share	-0.004	(0.003)	0.280	(0.060)	282
Top Incomes	Income	Inverted Pareto Lorenz Coefficient	-0.011	(0.007)	0.037	(0.013)	397
WDI/POVCAL	Income and Consumption Mixed	Gini Coefficient	-0.006	(0.002)	0.003	(0.001)	540
WDI/POVCAL	Per Capita Income	Gini Coefficient	-0.008	(0.001)	0.004	(0.001)	339
WDI/POVCAL	Per Capita Consumption	Gini Coefficient	-0.012	(0.005)	0.004	(0.002)	317
ATG	Welfare Concept Varies	Gini Coefficient	-0.015	(0.001)	0.007	(0.001)	1054
WIID	Household Equivalized Gross Income	Gini Coefficient	-0.036	(0.006)	1.722	(0.271)	184
WIID	Household Equivalized Disposable Income	Gini Coefficient	-0.008	(0.001)	0.329	(0.036)	489
WIID	Household Per Capita Disposable Income	Gini Coefficient	-0.019	(0.003)	0.539	(0.080)	519
WIID	Household Per Capita Expenditure	Gini Coefficient	-0.025	(0.009)	0.902	(0.326)	65
WIID	Average of Multiple Welfare Concepts	Gini Coefficient	-0.015	(0.002)	0.581	(0.074)	1151
SWIID	Market Income	Gini Coefficient	-0.037	(0.001)	0.018	(0.001)	2661
SWIID	Net Market Income	Gini Coefficient	-0.025	(0.001)	0.011	(0.000)	2662

Note: Table displays estimates of β convergence using the model presented in equation 3.2 using the instrumental variables approach described in section 3.3.

Table 3.2: Convergence Estimates 1988-2000

Source (j)	Welfare Concept (w)	Inequality Measure (m)	Slope Coefficient	Std error	Intercept Coefficient	Std error	N
CEPALSTAT	Per Capita Total Current Income	Gini Coefficient	-0.036	(0.014)	0.022	(0.007)	38
CEPALSTAT	Per Capita Total Current Income	Theil Index	-0.045	(0.022)	0.031	(0.013)	38
CEPALSTAT	Per Capita Total Current Income	Atkinson Index (1)	-0.030	(0.017)	0.016	(0.006)	38
LIS	Equivalized Disposable Household Income	Gini Coefficient	0.007	(0.011)	-0.001	(0.004)	19
LIS	Equivalized Disposable Household Income	Atkinson Index (1)	0.009	(0.015)	-0.000	(0.004)	19
OECD IDD	Equivalized Pre Tax and Transfer Income	Gini Coefficient	-0.053	(0.033)	0.030	(0.014)	19
OECD IDD	Equivalized Disposable Household Income	Gini Coefficient	0.018	(0.005)	-0.004	(0.001)	44
SEDLAC	Per Capita Disposable Household Income	Gini Coefficient	-0.045	(0.008)	0.024	(0.004)	57
SEDLAC	Per Capita Disposable Household Income	Theil Index	-0.073	(0.022)	0.038	(0.011)	57
SEDLAC	Per Capita Disposable Household Income	Atkinson Index (1)	-0.049	(0.008)	0.021	(0.003)	57
SEDLAC	Household Equivalized Income	Gini Coefficient	-0.044	(0.008)	0.023	(0.004)	57
SEDLAC	Household Equivalized Income	Theil Index	-0.071	(0.021)	0.034	(0.009)	57
SEDLAC	Household Equivalized Income	Atkinson Index (1)	-0.049	(0.008)	0.020	(0.003)	57
Top Incomes	Income	Top 1% Share	-0.021	(0.007)	0.305	(0.058)	157
Top Incomes	Income	Top 5% Share	-0.026	(0.008)	0.707	(0.152)	130
Top Incomes	Income	Inverted Pareto Lorenz Coefficient	0.005	(0.011)	0.013	(0.020)	179
WDI/POVCAL	Income and Consumption Mixed	Gini Coefficient	-0.013	(0.006)	0.007	(0.003)	115
WDI/POVCAL	Per Capita Income	Gini Coefficient	-0.021	(0.007)	0.011	(0.004)	80
WDI/POVCAL	Per Capita Consumption	Gini Coefficient	-0.030	(0.016)	0.012	(0.006)	30
ATG	Welfare Concept Varies	Gini Coefficient	-0.026	(0.004)	0.013	(0.002)	376
WIID	Household Equivalized Gross Income	Gini Coefficient	0.000	(0.011)	0.426	(0.332)	24
WIID	Household Equivalized Disposable Income	Gini Coefficient	-0.009	(0.006)	0.436	(0.176)	158
WIID	Household Per Capita Disposable Income	Gini Coefficient	0.000	(0.006)	0.250	(0.235)	131
WIID	Household Per Capita Expenditure	Gini Coefficient	-0.027	(0.014)	0.969	(0.542)	22
WIID	Average of Multiple Welfare Concepts	Gini Coefficient	-0.010	(0.005)	0.460	(0.210)	395
SWIID	Market Income	Gini Coefficient	-0.056	(0.003)	0.029	(0.001)	1123
SWIID	Net Market Income	Gini Coefficient	-0.037	(0.002)	0.017	(0.001)	1123

Note: Table displays estimates of β convergence using the model presented in equation 3.2 using the instrumental variables approach described in section 3.3.

Table 3.3: Convergence Estimates 2000-2012

Source (j)	Welfare Concept (w)	Inequality Measure (m)	Slope Coefficient	Std error	Intercept Coefficient	Std error	N
CEPALSTAT	Per Capita Total Current Income	Gini Coefficient	-0.013	(0.012)	0.002	(0.007)	96
CEPALSTAT	Per Capita Total Current Income	Theil Index	-0.038	(0.008)	0.009	(0.005)	96
CEPALSTAT	Per Capita Total Current Income	Atkinson Index (1)	-0.017	(0.012)	0.001	(0.005)	96
LIS	Equivalized Disposable Household Income	Gini Coefficient	-0.033	(0.009)	0.011	(0.003)	11
LIS	Equivalized Disposable Household Income	Atkinson Index (1)	-0.042	(0.012)	0.009	(0.003)	11
OECD IDD	Equivalized Pre Tax and Transfer Income	Gini Coefficient	-0.032	(0.019)	0.016	(0.008)	140
OECD IDD	Equivalized Disposable Household Income	Gini Coefficient	-0.016	(0.005)	0.005	(0.002)	128
SEDLAC	Per Capita Disposable Household Income	Gini Coefficient	-0.042	(0.012)	0.016	(0.006)	119
SEDLAC	Per Capita Disposable Household Income	Theil Index	-0.060	(0.010)	0.021	(0.006)	119
SEDLAC	Per Capita Disposable Household Income	Atkinson Index (1)	-0.051	(0.010)	0.013	(0.004)	119
SEDLAC	Household Equivalized Income	Gini Coefficient	-0.042	(0.012)	0.015	(0.006)	119
SEDLAC	Household Equivalized Income	Theil Index	-0.061	(0.010)	0.019	(0.006)	119
SEDLAC	Household Equivalized Income	Atkinson Index (1)	-0.052	(0.011)	0.012	(0.004)	119
Top Incomes	Income	Top 1% Share	0.006	(0.008)	0.064	(0.084)	126
Top Incomes	Income	Top 5% Share	-0.017	(0.007)	0.498	(0.157)	107
Top Incomes	Income	Inverted Pareto Lorenz Coefficient	0.010	(0.011)	-0.010	(0.021)	151
WDI/POVCAL	Income and Consumption Mixed	Gini Coefficient	-0.023	(0.003)	0.007	(0.002)	248
WDI/POVCAL	Per Capita Income	Gini Coefficient	-0.030	(0.004)	0.009	(0.002)	154
WDI/POVCAL	Per Capita Consumption	Gini Coefficient	0.009	(0.009)	-0.004	(0.003)	162
ATG	Welfare Concept Varies	Gini Coefficient	-0.023	(0.004)	0.007	(0.002)	390
WIID	Household Equivalized Gross Income	Gini Coefficient	-0.114	(0.016)	5.447	(0.716)	111
WIID	Household Equivalized Disposable Income	Gini Coefficient	-0.025	(0.003)	0.695	(0.086)	228
WIID	Household Per Capita Disposable Income	Gini Coefficient	-0.011	(0.011)	0.178	(0.353)	270
WIID	Household Per Capita Expenditure	Gini Coefficient	-0.081	(0.023)	3.315	(0.811)	20
WIID	Average of Multiple Welfare Concepts	Gini Coefficient	-0.038	(0.005)	1.380	(0.187)	477
SWIID	Market Income	Gini Coefficient	-0.032	(0.003)	0.014	(0.002)	1116
SWIID	Net Market Income	Gini Coefficient	-0.031	(0.002)	0.011	(0.001)	1117

Note: Table displays estimates of β convergence using the model presented in equation 3.2 using the instrumental variables approach described in section 3.3.

Table 3.4: Convergence Estimates 1983-1999

Source (j)	Welfare Concept (w)	Inequality Measure (m)	Slope Coefficient	Std error	Intercept Coefficient	Std error	N
CEPALSTAT	Per Capita Total Current Income	Gini Coefficient	-0.028	(0.016)	0.017	(0.008)	26
CEPALSTAT	Per Capita Total Current Income	Theil Index	-0.043	(0.025)	0.030	(0.015)	26
CEPALSTAT	Per Capita Total Current Income	Atkinson Index (1)	-0.021	(0.019)	0.011	(0.007)	26
LIS	Equivalized Disposable Household Income	Gini Coefficient	0.009	(0.003)	-0.001	(0.001)	37
LIS	Equivalized Disposable Household Income	Atkinson Index (1)	0.007	(0.005)	0.000	(0.001)	37
OECD IDD	Equivalized Pre Tax and Transfer Income	Gini Coefficient	-0.066	(0.015)	0.032	(0.007)	32
OECD IDD	Equivalized Disposable Household Income	Gini Coefficient	0.016	(0.003)	-0.004	(0.001)	57
SEDLAC	Per Capita Disposable Household Income	Gini Coefficient	-0.008	(0.009)	0.007	(0.004)	54
SEDLAC	Per Capita Disposable Household Income	Theil Index	-0.017	(0.029)	0.015	(0.013)	54
SEDLAC	Per Capita Disposable Household Income	Atkinson Index (1)	-0.009	(0.009)	0.007	(0.003)	54
SEDLAC	Household Equivalized Income	Gini Coefficient	-0.005	(0.009)	0.006	(0.004)	54
SEDLAC	Household Equivalized Income	Theil Index	-0.013	(0.029)	0.012	(0.012)	54
SEDLAC	Household Equivalized Income	Atkinson Index (1)	-0.007	(0.010)	0.006	(0.003)	54
Top Incomes	Income	Top 1% Share	-0.008	(0.007)	0.180	(0.041)	215
Top Incomes	Income	Top 5% Share	-0.009	(0.007)	0.333	(0.120)	182
Top Incomes	Income	Inverted Pareto Lorenz Coefficient	-0.010	(0.008)	0.037	(0.013)	250
WDI/POVCAL	Income and Consumption Mixed	Gini Coefficient	-0.010	(0.005)	0.005	(0.002)	148
WDI/POVCAL	Per Capita Income	Gini Coefficient	-0.009	(0.004)	0.005	(0.002)	107
WDI/POVCAL	Per Capita Consumption	Gini Coefficient	-0.017	(0.011)	0.006	(0.004)	42
ATG	Welfare Concept Varies	Gini Coefficient	-0.022	(0.003)	0.011	(0.001)	504
WIID	Household Equivalized Gross Income	Gini Coefficient	-0.025	(0.010)	1.272	(0.413)	39
WIID	Household Equivalized Disposable Income	Gini Coefficient	-0.002	(0.003)	0.231	(0.085)	186
WIID	Household Per Capita Disposable Income	Gini Coefficient	-0.013	(0.003)	0.649	(0.103)	150
WIID	Household Per Capita Expenditure	Gini Coefficient	-0.041	(0.012)	1.648	(0.435)	33
WIID	Average of Multiple Welfare Concepts	Gini Coefficient	-0.018	(0.003)	0.891	(0.132)	497
SWIID	Market Income	Gini Coefficient	-0.044	(0.002)	0.022	(0.001)	1383
SWIID	Net Market Income	Gini Coefficient	-0.027	(0.002)	0.012	(0.001)	1383

Note: Table displays estimates of β convergence using the model presented in equation 3.2 using the instrumental variables approach described in section 3.3.

Table 3.5: F Statistics for Tests of Hypothesis A

Metric (m)	Source (j)	Welfare Concepts (w vs. v)	Time Period (T)			
			1988-2012	1988-2000	2000-2012	1983-1999
Gini Coefficient	WDI	Income vs. Consumption	1.821			
			(0.180)			
Gini Coefficient	SEDLAC	Equivalized Income vs Per Capita Income	0.042	0.000	0.000	0.059
			(0.838)	(0.986)	(1.000)	(0.808)
Atkinson Index (1)	SEDLAC	Equivalized Income vs Per Capita Income	0.003	0.000	0.008	0.035
			(0.958)	(0.982)	(0.929)	(0.853)
Theil Index	SEDLAC	Equivalized Income vs Per Capita Income	0.038	0.006	0.005	0.010
			(0.846)	(0.937)	(0.946)	(0.922)
Gini Coefficient	SWIID	Net Income vs. Market Income	49.800	25.768	0.046	44.630
			(0.000)	(0.000)	(0.830)	(0.000)
Gini Coefficient	OECD	Disposable Income vs. Total Income	24.343		0.541	28.976
			(0.000)		(0.462)	(0.000)

Notes: P-values in parentheses. Tests only performed where there were at least 10 countries and 30 observations in each subset.

Table 3.6: F Statistics for Tests of Hypothesis B

Welfare Concept (w)	Source (j)	Inequality Measures (l vs. m)	Time Period (T)			
			1988-2012	1988-2012	1988-2012	1988-2012
Per Capita Total Current	CEPAL	Gini Coefficient vs. Theil Index	0.968	0.108	3.151	
Income			(0.326)	(0.743)	(0.077)	
Per Capita Total Current	CEPAL	Atkinson Index (1) vs. Theil Index	1.518	0.283	2.189	
Income			(0.219)	(0.596)	(0.141)	
Per Capita Total Current	CEPAL	Gini Coefficient vs. Atkinson Index(1)	0.119	0.078	0.075	
Income			(0.730)	(0.781)	(0.784)	
Per Capita Disposable	SEDLAC	Gini Coefficient vs. Atkinson Index(1)	0.412	0.179	0.333	0.004
Household Income			(0.521)	(0.673)	(0.564)	(0.947)
Per Capita Disposable	SEDLAC	Gini Coefficient vs. Atkinson Index(1)	0.601	0.198	0.419	0.012
Household Income			(0.439)	(0.658)	(0.518)	(0.914)
Per Capita Pre-tax Income	WTID	Top five percent Income Share vs.	2.064	7.189	3.354	0.000
		Inverted Pareto Lorenz Coefficient	(0.151)	(0.008)	(0.069)	(0.997)
Per Capita Pre-tax Income	WTID	Top one percent Income Share vs.	0.268	4.742	0.143	0.001
		Inverted Pareto Lorenz Coefficient	(0.605)	(0.030)	(0.706)	(0.979)
Per Capita Total Current	WTID	Top one percent Income Share vs. Top	0.647	0.344	0.940	0.000
Income		five percent Income Share	(0.422)	(0.558)	(0.333)	(0.995)
Per Capita Disposable	LIS	Gini Coefficient vs. Atkinson Index(1)	1.277	0.020	0.325	0.122
Household Income			(0.260)	(0.889)	(0.576)	(0.728)

Notes: P-values in parentheses. Tests only performed where there were at least 10 countries and 30 observations in each subset.

Table 3.7: F-Statistics for Tests of Hypothesis C

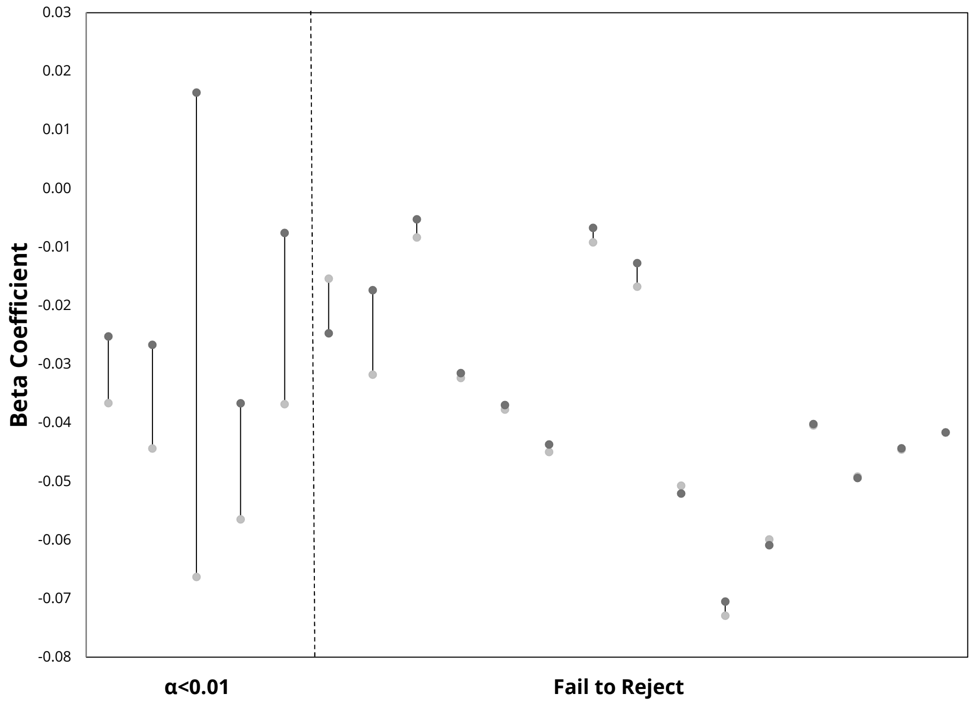
Welfare Concept (w)	Inequality Measure (m)	Source (j vs. k)	Time Period (T)			
			1988-2012	1988-2012	1988-2012	1988-2012
Equivalized Disposable or Net Household Income	Gini Coefficient	LIS vs. OECD IDD	0.014 (0.905)	1.206 (0.277)	0.089 (0.766)	1.613 (0.208)
Equivalized Disposable or Net Household Income	Gini Coefficient	SWIID vs. LIS	0.085 (0.771)	6.227 (0.013)	5.568 (0.020)	21.179 (0.000)
Equivalized Disposable or Net Household Income	Gini Coefficient	SWIID vs. OECD IDD	12.025 (0.001)	17.293 (0.000)	1.895 (0.169)	55.831 (0.000)
Equivalized Disposable or Net Household Income	Gini Coefficient	SWIID vs. SEDLAC	0.040 (0.842)	2.491 (0.116)	0.063 (0.803)	10.619 (0.001)
Equivalized Disposable or Net Household Income	Gini Coefficient	SWIID vs. LIS and SEDLAC	1.903 (0.168)	1.156 (0.283)	1.926 (0.166)	21.930 (0.000)
Equivalized Disposable or Net Household Income	Gini Coefficient	WIID vs. LIS	0.108 (0.742)	0.421 (0.517)	0.199 (0.656)	11.618 (0.001)
Equivalized Disposable or Net Household Income	Gini Coefficient	WIID vs. OECD IDD	0.168 (0.682)	2.035 (0.156)	0.062 (0.804)	25.613 (0.000)
Equivalized Disposable or Net Household Income	Gini Coefficient	WIID vs. SEDLAC	0.083 (0.774)	1.844 (0.180)	11.524 (0.001)	0.808 (0.372)
Equivalized Disposable or Net Household Income	Gini Coefficient	WIID vs. LIS and SEDLAC	5.184 (0.023)	1.017 (0.315)	4.041 (0.046)	4.103 (0.044)
Equivalized Disposable or Net Household Income	Gini Coefficient	WIID vs. SWIID	19.726 (0.000)	1.313 (0.252)	4.524 (0.034)	2.445 (0.118)
Per Capita Expenditure or Consumption	Gini Coefficient	WIID vs. WDI/POVCAL	1.089 (0.298)	0.111 (0.742)	24.787 (0.000)	0.057 (0.812)
Per Capita Disposable Household Income	Gini Coefficient	WIID vs. SEDLAC	0.463 (0.497)			
Equivalized Market or Pre Tax and Transfer Income	Gini Coefficient	SWIID vs. OECD IDD	0.044 (0.834)		2.827 (0.093)	0.667 (0.415)
Aggregated Welfare Concepts	Gini Coefficient	WIID vs. ATG	0.291 (0.589)	5.645 (0.018)	8.923 (0.003)	0.577 (0.448)
Aggregated Welfare Concepts	Gini Coefficient	WIID vs. WDI/POVCAL	20.747 (0.000)	0.190 (0.663)	8.555 (0.004)	7.941 (0.005)
Aggregated Welfare Concepts	Gini Coefficient	ATG vs. WDI/POVCAL	26.212 (0.000)	6.318 (0.012)	0.050 (0.824)	9.784 (0.002)

Notes: P-values in parentheses. Tests only performed where there were at least 10 countries and 30 observations in each subset.

Table 3.8: Summary of Results

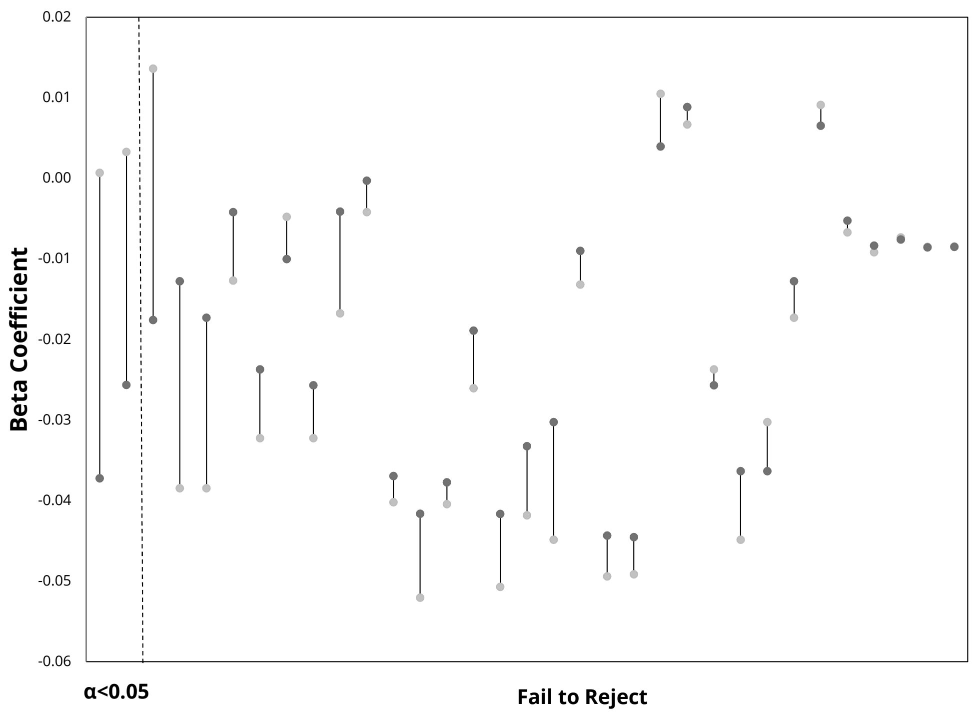
Hypothesis		Tests	Rejections
A	All else equal, estimates are identical across welfare concepts	20	5
B	All else equal, estimates are identical across inequality measures	33	2
C	All else equal, estimates are identical across data sources	60	25
D	All else equal, estimates are identical across regional subsets of countries	191	111
E	All else equal, estimates are identical across time periods	153	57

Figure 3.1: Distribution of Coefficients in Pairwise Tests of Hypothesis A



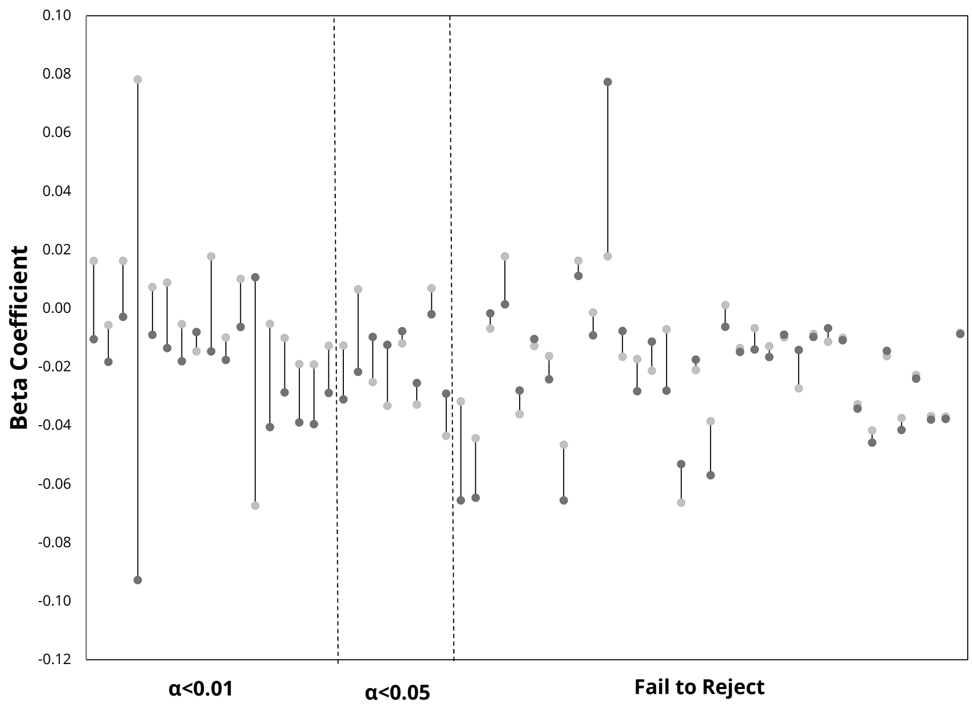
Note: The figure displays the beta coefficients estimated for the pairwise F-tests described in section 3.4 for Hypothesis A. Vertical lines connect the coefficients for each test. Tests are sorted by P-value with the strongest statistical significance on the left.

Figure 3.2: Distribution of Coefficients in Pairwise Tests of Hypothesis B



Note: The figure displays the beta coefficients estimated for the pairwise F-tests described in section 3.4 for Hypothesis B. Vertical lines connect the coefficients for each test. Tests are sorted by P-value with the strongest statistical significance on the left.

Figure 3.3: Distribution of Coefficients in Pairwise Tests of Hypothesis C



Note: The figure displays the beta coefficients estimated for the pairwise F-tests described in section 3.4 for Hypothesis C. Vertical lines connect the coefficients for each test. Tests are sorted by P-value with the strongest statistical significance on the left.

Appendix A

Background on AmeriCorps

A.1 What is AmeriCorps?

AmeriCorps is a network of service programs throughout the United States coordinated by the Corporation for National and Community Service (CNCS).[1] Through CNCS, AmeriCorps receives federal funds with which it both administers the network of service programs and provides funding and benefits to its members. There are currently three types of AmeriCorps programs: VISTA, National Civilian Community Corps (NCCC), and State and National Direct.[2] In most cases, members join for a year of intensive community service.[3] The type and structure of their service, however, varies greatly. Some members serve in classrooms while other work on the Appalachian Trail. Some work for nonprofit organizations, while others work for the Federal Emergency Management Agency. Collectively, there are more than 75,000 AmeriCorps members in the U.S. and more than 800,000 alumni.

[1] Unless cited as from a different source, the information in this section was retrieved from the official website of CNCS http://www.nationalservice.gov.

[2] The 2015 CNCS budget would bring the Senior Companions Program and Foster Grandparents Program into AmeriCorps. Both programs are currently part of Senior Corps.

[3] One year is the norm, but there are exceptions within the State and National program. Some states allow AmeriCorps grants to be split between 2, 3, or 4 part-time or shorter-term members. Members of Teach for America are part of the State and National program, but commit to two years of service.

AmeriCorps was officially created by the National and Community Service Trust Act of 1993, but its origins really begin with President Lyndon Johnson's war on poverty. The Economic Opportunity Act of 1964 created Volunteers in Service to America (VISTA) and, in December of that year, President Johnson launched the first class VISTA members into action with a White House ceremony. Early VISTA members were pivotal in connecting local community organizations with new federal programs, such as Head Start, that were launched during Johnson's presidency. The legacy of fighting poverty remains central to the AmeriCorps VISTA program and VISTA continues to place its members in low-income communities through the U.S.. Today, AmeriCorps VISTA members work directly with local "sponsor" organizations for whom they serve capacity building roles, helping local charities and nonprofits expand their ability to provide services to the community.

In 1992, Congress created the National Civilian Community Corps. Modelled after the depression era Civilian Community Corps program, NCCC places young Americans, between 18 and 24 years of age, on teams that provide support to local communities. Typically, each NCCC team takes part in four project "spikes" during their service year. During the spikes, members may work with local nonprofit organizations or government agencies. NCCC projects range from fighting wild-land fires to working in inner city elementary schools.

With VISTA and NCCC were already in place, AmeriCorps was created with the passage of the National and Community Service Trust Act in 1993. The act created CNCS and tasked it with administering three programs: AmeriCorps, Senior Corps, and Learn and Serve America.[4] The pre-existing VISTA and NCCC programs were brought under the AmeriCorps umbrella and the new, larger, AmeriCorps State and

[4] Learn and Serve America was a grant program that funded service learning in schools. It was defunded and eliminated in 2011. Senior Corps brought together three existing programs, Retired Senior Volunteer Program (RSVP), the Foster Grandparent Program, and the Senior Companion program. The Fiscal Year 2015 budget would eliminate Senior Corps by reallocating funding for RSVP to the Volunteer Generation fund and bringing the Foster Grandparent and Senior Companion programs into AmeriCorps.

National program was created.

AmeriCorps State and National places its members in “direct service”, embedded in nonprofit organizations and government agencies. The State and National programs are the most diverse for two reasons. First, there are multiple types of grants available to fund State and National programs. Second, each state’s service commission determines the types of service projects that are eligible for AmeriCorps grants. AmeriCorps State and National members might work in schools, build low income housing, do outreach with the homeless, or work to plan a fundraising gala. Members of many national service programs such as Teach for America, City Year, and Campus Compact are State and National AmeriCorps.

A.2 How AmeriCorps is Funded

AmeriCorps programs are financed with a combination of funding from all levels of government and both the public and nonprofit sectors. Administration of AmeriCorps and associated costs are the responsibility of CNCS and CNCS funds at least a portion of every AmeriCorps program. CNCS receives funds through congressional appropriations and in turn funds AmeriCorps, Senior Corps, the National Service Trust, and the Social Innovation and George H.W. Bush Volunteer Generation funds. Federal AmeriCorps funding pays for CNCS to directly fund and administer the NCCC and VISTA programs, fund State and National programs, administer the National programs, and provide funding to state service commissions to administer the State programs. Funding from the private sector and local governments varies by program. Each program has a “sponsor” responsible for funding part of the program, although the level of funding varies greatly.

Over the last decade, Federal funding for NCCC has ranged between \$24 and \$32 million annually. This funding pays for the operation of five regional campuses: Denver, CO; Sacramento, CA; Perry Point, MD; Vicksburg, MS; and Vinton, IA. NCCC also provides its members with a uniform, housing, meals, and medical benefits along with a small living allowance of about $4,000. Additionally, each team is provided with an NCCC van for transportation. When NCCC members are participating in off-campus projects, the project sponsors provide or arrange for housing and meals. All other costs are paid out of the program's budget with the result that each NCCC member costs CNCS about $25,000.

CNCS receives between $85 and $100 million in annual appropriations for AmeriCorps VISTA.[5] VISTA members receive a living allowance of about $9,500 paid directly from AmeriCorps VISTA. The funding for this stipend may come from either the sponsor organization or CNCS. Whether the stipend is funded by the sponsor organization or by CNCS, payroll taxes are paid directly by CNCS. Health insurance and liability coverage for VISTA members are also provided by and paid for by CNCS. VISTA provides basic project management training to all of its members. Sponsors are only required to provide work-related resources such as office space and transportation. While VISTA members are intended to augment, rather than replace existing workers, the effect is that organizations gain a full-time worker at minimal cost and sometimes for free. The average cost to CNCS runs between $17,000 and $18,000 per VISTA member.

Funding for State and National programs has both risen and fallen in the last decade. At a minimum, the fiscal year 2008 budget allotted State and National programs was $257 million. Funding from the 2009 Recovery Act and expansion under the Edward M. Kennedy Serve America Act lead to a peak level of funding of $360 million in 2009 and $370 million in 2010. The allocation of that funding into various grant types has changed over this time as well. Today, there are four types or grant programs: State Formula, U.S. Territories, Indian Tribe, and Single and Multi-

[5] AmeriCorps VISTA received an additional $65 million appropriation as part of the 2009 Recovery Act.

State Competitive. State Formula grants, which comprise 35 percent of program funding, are given to the state service commissions. State service commissions are then responsible for selecting sponsors and administering programs. U.S. Territories grants receive one percent of funds and are directed to territory service commissions. Indian tribe grants, which also receive one percent of funds, are awarded to U.S. Indian tribes in a competitive process. The remaining funds are available through competitive grant processes at both the state and national level.

Most State and National programs require a cost-share between the sponsor organization and AmeriCorps. Members are paid directly by the sponsor organizations. The sponsors are then eligible for partial reimbursement from either the state service commission or CNCS. In some cases, the sponsor organizations apply for grants that only reimburse them for minimal administrative costs. The most generous grants reimburse up to $16,000 of program costs with the requirement that the sponsor matches at least 24 percent of the cost. The average per member cost to CNCS runs between $7,800 and $9,600 per State and National AmeriCorps member.

All AmeriCorps members, except those who have two or more prior years of service, are eligible to receive the Segal AmeriCorps Education Award. The education award is paid for from CNCS's National Service Trust. The value of the award is equal to the maximum value of a Pell Grant in the year in which the member serve – currently $5,645.[6] AmeriCorps members may direct payment of the award only to qualified education institution and student loan providers. VISTA members may choose to forgo the education award in exchange for a smaller, lump-sum cash stipend.

[6] AmeriCorps members in part-time or part year State and National programs receive a partial award. For example, a full-time, half-year member would receive half of the total award value.

A.3 How Organizations Get AmeriCorps

Applying to sponsor an AmeriCorps project is, in many ways, similar to applying for a grant. Prospective sponsors must fill out formal applications that describe their goals, establish performance measures, and contain a detailed budget. Each type of AmeriCorps program has its own application process for prospective sponsors.

Nonprofits, municipalities, state and local government agencies, schools, and Indian tribes are all eligible to sponsor NCCC projects. Project sponsors either provide or arrange and pay for housing and meals for the NCCC members during the project. Additionally, they are required to supply materials, tools, training, and supervision to their NCCC teams. All sponsors must be able to justify their project as meeting one of the following community needs: natural and other disasters, infrastructure improvement, environmental stewardship and conservation, energy conservation, and urban and rural development. Applications are submitted and reviewed by regional offices aligned with each campus.

Organizations who want to sponsor AmeriCorps VISTAs apply through CNCS state offices and the use of the web-based *egrants* application system. All VISTA positions are full-time and for a full year of service. Only government agencies and nonprofit organizations may sponsor VISTA members. VISTA projects must address the needs of low-income communities and must be focused on capacity building rather than direct service. In constructing their project application, potential sponsor organizations must decide whether or not to cost-share. Organizations who chose to cost-share reimburse VISTA for part of the member's stipend.

Potential sponsors for State and National Direct AmeriCorps programs can apply either through state service commissions or directly with CNCS.[7] Each year, CNCS releases a Notice of Federal Funding Opportunity for AmeriCorps State and National

[7] Only nonprofit organizations that operate in multiple states, consortium or nonprofit organizations, and Indian tribes may apply directly with CNCS.

that outlines the application process. Potential sponsors then need to decide which type of program (among the various types of grants and cost-share arrangements and between state-level or national funds) is right for them and provide the documentation specific to its application. All State and National programs are competitive and, while preference is given to continuing existing (successful) programs, CNCS makes planning grants available to organizations that intend to design new AmeriCorps programs.

A.4 How Members Join

How a member joins AmeriCorps varies depending on the program he or she is joining. Whether joining NCCC, VISTA, or a State or National program, most people begin at the *My AmeriCorps* website, `http://my.americorps.gov`. Through the website, potential members can search various service opportunities. Much like *Monster.com* and other job search websites applicants can search by location, industry (here "service area") or specific skill requirements. Potential members can then view the position's description, requirements, and term of service. Applications are then submitted through the *My AmeriCorps* website.

Applicants to be NCCC team members apply to postings based on the term of service, not to the specific team or campus. Generally, team members must be between 18 and 24 years of age. However, applicants may apply to be a Team Leader (TL) with no upper age limit. Members who are accepted are assigned to one of the five regional campuses. At each campus, members receive basic instructions on NCCC regulations, first aid, CPR, and general public safety. They are then assigned to their eight to twelve person team. Service projects or "spikes", which typically last six to eight weeks, are then assigned to each team.

Applicants to VISTA and State and National programs apply for specific roles in specific programs. The decisions on which applicants to accept are made by the sponsor organizations As such, sponsors may post ads on other job search websites that direct applicants to the *My AmeriCorps* website. In some cases, organizations recruit and interview applicants directly and then direct their preferred candidates to the website to complete the hiring process.

A.5 Legislation and Expansion

As a program created and funded by the United States Government, the size AmeriCorps program is dictated by political forces. The VISTA program began in 1964, and while AmeriCorps was not created until 1993, a similar proposal was put forward by Senators Sam Nunn and Dave McCurdy in 1989 (Just, 2003). The political momentum for AmeriCorps came from President Clinton's election campaign. Even then, he was only able to get congressional funding for a first class of 20,000 members.

By the time Clinton left office, AmeriCorps had grown to 50,000 members, and it would not be long before a major push for further expansion (Just, 2003). In 2001, shortly after the September 11th attacks, Senators John McCain and Evan Bayh announced a plan to increase the size of AmeriCorps five-fold, from 50,000 to 250,000 members annually (Sack, 2001). In early 2002, President Bush proposed the USA Freedom Corps, a wide-ranging plan to support and expand national service. His proposal would have increased the size of AmeriCorps by 200,000 and he quickly promised to work with Senators McCain and Bayh to meet that goal. The proposal was not univerisal popular, however. When a version of the bill finally made it out of committee then house majority leader Congressman Dick Armey blocked the bill from coming to a vote.

Without major legislation, the Bush Administration was able to gradually increase funding for AmeriCorps beginning in 2004. By 2009, AmeriCorps had expanded to

75,000 members, yet calls for greater expansion of AmeriCorps had never gone away (Associated Press, 2009). In 2008 Senator Kennedy, a Democrat, and Senator Orrin Hatch, a Republican proposed a plan to increase the size of AmeriCorps to 250,000 members, the same number proposed 6 years earlier by Senators McCain and Bayh. Soon after, expanding national service became piece of Senator Barack Obama's 2008 presidential campaign (The Editorial Board, 2014).

Shortly after President Obama took office, the American Recovery and Reinvestment Act (the stimulus package) was passed. Included among the other stimulative spending was a short-term, $200 million increase in funding for AmeriCorps. The VISTA program saw a one year increase in membership of over 40% before returning to the prior year's member levels. The stimulus spending was, by its nature, one-time-only and proponents of expanding AmeriCorps quickly turned to a more permanent plan.

In March of 2009, an ailing Senator Kennedy, then being treated for brain cancer, returned to Washington to help pass the Serve America Act through a congress dominated by members of his and the President Obama's party (Associated Press, 2009). By the time it was signed into law later that year, the act had his name on it. The law put AmeriCorps on a path to 250,000 members a year, with the growth targeted exclusively at the State and National Direct family of programs. The 2010-2011 class of AmeriCorps*State and National Direct was more than 10 percent larger than the prior cohort, with growth planned through 2017. 2011, however, proved to be the high point for AmeriCorps membership (The Editorial Board, 2014). The 2010 congressional election created divided legislature that has not provided funding for the Kennedy Serve America Act. Rather, the most recent report of the Budget Committee of the House of Representatives proposes eliminating AmeriCorps and CNCS entirely (?).

Appendix B

Incorporating Data from WIID 3.3 into the Sensitivity Analysis

The World Income Inequality Database (WIID) is a secondary source database that collects inequality indicators form a variety of sources. WIID data is characterized by welfare definition, unit of analysis,income share unit, and a number of variables that describe the breadth or coverage of the inequality estimate. We use WIID 3.3, released in September of 2015.

In our analysis, we only use inequality indicators that are classified as including all age groups and all regions within a country. We then divide the dataset into four groups by welfare definition; they are consumption, gross income, disposable income, and other. We then expand the dataset from four categories to 12 based on whether the unit of analysis is household, person, or other. We further divide the dataset, from 12 categories to 36 based on whether the indicators are calculated either using household adult equivalence scales, by individual or household per capita, or some other equivalency scale. Finally, we expand from 36 categories to 144 based on quality rating: high, medium, low, or unknown. Where multiple estimates exist for the same welfare definition, scale, quality, country, and year, we average the indicators together.

The result is 144 unique panels of inequality estimates based on WIID 3.3.

We then limit the number of panels by using a choice by precedence approach to 48 by collapsing along the dimension of unit of analysis into panels of country, year, welfare definition, and equivalence scale. Within each cell we keep the best available indicator, categorizing the best unit of analysis as person, the second best unit of analysis as household, and the worst unit of analysis as other. We further limit the number of panels by merging all panels that list the welfare metric as " other" whether they are measured using household adult equivalence scales, by individual or household per capita, or some other equivalency scale. This reduces the number of panels to 40. That is, for each level of quality there are ten unique panels: (1) disposable household equivalent income, (2) gross household equivalent income, (3) household equivalent expenditure, (4) disposable household per capita income, (5) gross household per capita income, (6) per capita expenditure, (7) other disposable income, (8) other gross income, (9) other expenditure, and (10) other.

Next, within each of the ten categories above, we replace metrics of medium, low, or unknown quality with better quality metrics when possible. After this process the high quality panels are subsets of the medium quality panels, the medium quality panels are subsets of the low quality panels, and the low quality panels are subsets of the unknown quality panels. Unless otherwise stated, we the figures that appear in this paper are based on the medium quality panels.

References

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

Abadie, A., Diamond, A., and Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *The American Economic Review*, 93(1):113–132.

Abrams, B. A. and Schmitz, M. D. (1984). The crowding-out effect of governmental transfers on private charitable contributions: cross-section evidence. *National Tax Journal*, pages 563–568.

Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2013). Democracy, redistribution and inequality. Working Paper 19746, National Bureau of Economic Research.

Adhikari, B. and Alm, J. (2015). Evaluating the economic effects of flat tax reforms using synthetic control methods. *Tulane University Working Paper*.

Alaska Department of Revenue (2014). Alaska indirect expenditure report. preliminary report for FY2009-FY2013.

Alaska Department of Revenue (2015). Alaska's education tax credit program information and documents. http://dor.alaska.gov/EducationTaxCreditInformationandDocuments.aspx. Accessed March 10, 2015.

Andreoni, J. (1988). Privately provided public goods in a large economy: the limits of altruism. *Journal of Public Economics*, 35(1):57–73.

Andreoni, J. (1989). Giving with impure altruism: Applications to charity and ricardian equivalence. *Journal of Political Economy*, 97(6):1447–1458.

Andreoni, J. (1990). Impure altruism and donations to public goods: a theory of warm-glow giving. *The Economic Journal*, 100(401):464–477.

Andreoni, J., Payne, A., and Smith, S. (2014). Do grants to charities crowd out other income? evidence from the uk. *Journal of Public Economics*, 114:75–86.

Andreoni, J. and Payne, A. A. (2003). Do government grants to private charities crowd out giving or fund-raising? *American Economic Review*, pages 792–812.

Andreoni, J. and Payne, A. A. (2011). Is crowding out due entirely to fundraising? evidence from a panel of charities. *Journal of Public Economics*, 95(5):334–343.

Andreoni, J. and Rao, J. M. (2011). The power of asking: How communication affects selfishness, empathy, and altruism. *Journal of Public Economics*, 95(7):513–520.

Associated Press (2009). Senate to consider expanding americorps.

Atkinson, A. and Bourguignon, F., editors (2015). *Handbook of Income Distribution*, volume 2A-2B of *Handbook of Income Distribution*. Elsevier, 1 edition.

Auten, G. E., Sieg, H., and Clotfelter, C. T. (2002). Charitable giving, income, and taxes: An analysis of panel data. *The American Economic Review*, 92(1):371–382.

Bakija, J. and Heim, B. T. (2011). How does charitable giving respond to incentives and income? New estimates from panel data. *National Tax Journal*, 64(2):615–650.

Barone, G. and Mocetti, S. (2014). Natural disasters, growth and institutions: a tale of two earthquakes. *Journal of Urban Economics*, 84:52–66.

Barrett, K. S., McGuirk, A. M., and Steinberg, R. (1997). Further evidence on the dynamic impact of taxes on charitable giving. *National Tax Journal*, 50(2):321–34.

Benabou, R. (1996). Inequality and growth. In *NBER Macroeconomics Annual 1996, Volume 11*, pages 11–92. MIT Press.

Bergstrom, T., Blume, L., and Varian, H. (1986). On the private provision of public goods. *Journal of Public Economics*, 29(1):25–49.

Billmeier, A. and Nannicini, T. (2013). Assessing economic liberalization episodes: A synthetic control approach. *Review of Economics and Statistics*, 95(3):983–1001.

Blasnik, M. (2010). Reclink: Stata module to probabilistically match records.

Bleaney, M. and Nishiyama, A. (2003). Convergence in income inequality: differences between advanced and developing countries. *Economics Bulletin*, 4(22):1–10.

Bloom, N., Griffith, R., and Van Reenen, J. (2002). Do R&D tax credits work? evidence from a panel of countries 1979–1997. *Journal of Public Economics*, 85(1):1–31.

Bourguignon, F. (2015). *The Globalization of Inequality*. Princeton University Press.

Breman, A. (2006). The economics of altruism, paternalism and self-control.

Brudney, J. L. and Gazley, B. (2002). Testing the conventional wisdom regarding volunteer programs: A longitudinal analysis of the service corps of retired executives and the us small business administration. *Nonprofit and Voluntary Sector Quarterly*, 31(4):525–548.

Bureau of Economic Analysis (2014). Sa1-3 personal income summary. electronic. accessed February 28, 2014.

Button, P. (2015). Do Tax Incentives Affect Business Location? Evidence from Motion Picture Production Incentives. Working Papers 1507, Tulane University, Department of Economics.

Chen, S. and Ravallion, M. (2001). How did the world's poorest fare in the 1990s? *Review of Income and wealth*, 47(3):283–300.

Conn. Gen. Stat. tit. 12 Ch. 228a §12-630aa-638. R.E. van Norstrand neighborhood assistance act. http://www.cga.ct.gov/current/pub/chap_228a.htm. Accessed March 10, 2013.

Corporation for National and Community Service (2009). Notice of federal funding opportunity: Americorps state and national grants fy 2010. CFDA Number: 94.006.

Day, K. and Devlin, R. (1996). Volunteerism and crowding out: Canadian econometric evidence. *Canadian Journal of Economics*, pages 37–53.

De Vita, C. J. and Twombly, E. C. (2004). Charitable tax credits: Boon or bust for nonprofits? Charting Civil Society 16, The Urban Institute, Washington, DC.

Deininger, K. and Squire, L. (1996). A new data set measuring income inequality. *The World Bank Economic Review*, 10(3):565–591.

Del. Code tit. 30 Ch. 11 §. 2001-2007. Neighborhood assistance tax credit. http://delcode.delaware.gov/title30/c020/sc01/index.shtml. Accessed March 10, 2015.

Department of Finance (2011). State of Delaware tax preference report, 2011 edition. Division of Revenue. http://finance.delaware.gov/publications/tax_prefer/report_11.pdf. Accessed March 10, 2015.

Department of Social Services (2014). Neighborhood assistance program (NAP) project fact sheet July 1, 2014 June 30, 2015. Commonwealth of Virginia.

Division of Revenue. State of Delaware (August 26, 1999). Technical information memorandum 99-4. http://revenue.delaware.gov/information/tims/tim99-4.html. Accessed March 10, 2015.

Duncan, B. (1999). Modeling charitable contributions of time and money. *Journal of Public Economics*, 72(2):213–242.

Duncan, B. (2004). A theory of impact philanthropy. *Journal of Public Economics*, 88(9):2159–2180.

Duquette, N. J. (2016). Do tax incentives affect charitable contributions? evidence from public charities' reported revenues. *Journal of Public Economics*, 137:51 – 69.

Eckel, C. C., Grossman, P. J., and Johnston, R. M. (2005). An experimental test of the crowding out hypothesis. *Journal of Public Economics*, 89(8):1543–1560.

Ezcurra, R. and Pascual, P. (2005). Is there convergence in income inequality levels among the european regions? *Applied Economics Letters*, 12(12):763–767.

Ezcurra, R. and Pascual, P. (2009). Convergence in income inequality in the united states: a nonparametric analysis. *Applied Economics Letters*, 16(13):1365–1368.

Faulk, D. (2002). Do state economic development incentives create jobs? an analysis of state employment tax credits. *National Tax Journal*, 55(2):263–280.

Ferreira, F. H. and Lustig, N., editors (2015). *Special Issue on "Apraising Cross-National Income Inequality Databases"*, volume 13. Springer.

Ferreira, F. H. G., Lustig, N., and Teles, D. (2015). Appraising cross-national income inequality databases: An introduction. *The Journal of Economic Inequality*, 13(4):497–526.

Frank, M. W. (2009). Inequality and growth in the United States: Evidence from a new state-level panel of income inequality measures. *Economic Inquiry*, 47(1):55–68.

Frumkin, P., Jastrzab, J., Vaaler, M., Greeney, A., Grimm, R. T., Cramer, K., and Dietz, N. (2009). Inside national service: Americorps' impact on participants. *Journal of Policy Analysis and Management*, 28(3):394–416.

Gene, K. (2013). 2009 contributions to qualifying charitable organizations tax credit: A summary of activity. Arizona Department of Revenue Office of Economic Research & Analysis.

Gomes, F. (2007). Convergence in income inequality: the case of brazilian municipalities. *Economics Bulletin*, 15(15):1–9.

Governor's Office for Economic Analysis, Office of State Budget Directory (2011). Tax expenditure analysis fiscal year 2012 - 2014. Commonwealth of Kentucky.

Gruber, J. and Hungerman, D. M. (2007). Faith-based charity and crowd-out during the great depression. *Journal of Public Economics*, 91(5):1043–1069.

Gullickson, A. and Tilkes, J. (2013). Iowa's Endow Iowa tax credit tax credits program evaluation study. Tax Research and Program Analysis Section, Iowa Department of Revenue.

Handy, F., Mook, L., and Quarter, J. (2008). The interchangeability of paid staff and volunteers in nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, 37(1):76–92.

Heutel, G. (2014). Crowding out and crowding in of private donations and government grants. *Public Finance Review*, 42(2):143–175.

Ho, T.-w. (2015). Income inequality may not converge after all: Testing panel unit roots in the presence of cross-section cointegration. *The Quarterly Review of Economics and Finance*, 56:68–79.

Iowa Council of Foundations and Iowa Gaming Association (2013). Building Iowa's communities through philanthropy: A snapshot of the 2012-2013 county endowment fund program & grants made by casino nonprofit organizations.

Iowa Council of Foundations and the Iowa Economic Development Authority (2014). Endow Iowa Tax Credit and Country Endowment Fund Programs - 2013, A Report to the Governor and the Iowa Legislature.

Iowa Department of Revenue (2014). County endowment program distribution histories for local government. Microsoft Excel File. https://tax.iowa.gov/report/Distributions. Accessed March 18, 2015.

IRS Statistics of Income Division. Sample of exempt organizations [1985-2011].

Johnson Center for Philanthropy (2013). Impact on giving after the repeal of the Michigan community foundation tax credit. Grand Valley State University. https://www.michiganfoundations.org/sites/default/files/resources/Impact-of-Giving-After-Repeal-of-Tax-Credit-Johnson-Center-2013.pdf. Accessed: February 13, 2015.

Just, R. (2003). What ever happened to national service? how a bush policy pledge quietly dissappeared. *Washington Monthly*.

Kansas Department of Commerce (2014). Community service tax credit program. http://www.kansascommerce.com/index.aspx?NID=117. Accessed February 4, 2015.

Khanna, J., Posnett, J., and Sandler, T. (1995). Charity donations in the UK: New evidence based on panel data. *Journal of Public Economics*, 56(2):257–272.

Khanna, J. and Sandler, T. (2000). Partners in giving:: The crowding-in effects of uk government grants. *European Economic Review*, 44(8):1543–1556.

Kingma, B. R. (1989). An accurate measurement of the crowd-out effect, income effect, and price effect for charitable contributions. *The journal of political economy*, pages 1197–1207.

Ky. Rev. Stat. §141.438. Endow Kentucky tax credit. http://www.lrc.ky.gov/statutes/statute.aspx?id=43537. Accessed March 10, 2015.

Li, H. and Zou, H.-f. (1998). Income inequality is not harmful for growth: theory and evidence. *Review of development economics*, 2(3):318–334.

Lin, P.-C. and Huang, H.-C. (2012). Convergence in income inequality? evidence from panel unit root tests with structural breaks. *Empirical Economics*, 43(1):153–174.

Lokshin, B. and Mohnen, P. (2012). How effective are level-based R&D tax credits? evidence from the Netherlands. *Applied Economics*, 44(12):1527–1538.

Louisiana Department of Revenue (2013). 2011 — 2012 annual tax collection report. http://revenue.louisiana.gov/Publications/LDR_Annual_Report(11Accessed March 11, 2015.

Louisiana Department of Revenue (2015). School readiness tax credits. http://www.rev.state.la.us/Miscellaneous/SchoolAccessed March 11, 2015.

Lustig, N., editor (2016). *Commitment to Equity Handbook. A Guide to Estimating the Impact of Fiscal Policy on Inequality and Poverty.*

Lustig, N., Pessino, C., and Scott, J. (2014). The impact of taxes and social spending on inequality and poverty in argentina, bolivia, brazil, mexico, peru, and uruguay: Introduction to the special issue. *Public Finance Review*, 42(3):287–303.

Menchik, P. and Weisbrod, B. (1987). Volunteer labor supply. *Journal of Public Economics*, 32(2):159–183.

Mideksa, T. K. (2013). The economic impact of natural resources. *Journal of Environmental Economics and Management*, 65(2):277–289.

Missouri Department of Economic Development (2013). Youth opportunities program 2013 annual report. http://ded.mo.gov/upload/Genasrep2013.pdf. Accessed March 10, 2015.

Missouri Department of Economic Development (2015). Youth opportunity program. http://ded.mo.gov/BCSAccessed March 10, 2015.

Moretti, E. and Wilson, D. J. (2014). State incentives for innovation, star scientists and jobs: Evidence from biotech. *Journal of Urban Economics*, 79:20 – 38.

Name-Correa, A. and Yildirim, H. (2013). A theory of charitable fund-raising with costly solicitations. *American Economic Review*, 103(2):1091–1107.

National Council of Nonprofits (2015). State charitable giving incentives. https://www.councilofnonprofits.org/trends-policy-issues/state-charitable-giving-incentives. Accessed January 22
2015.

Nebraska Department of Revenue (2010). Nebraska endowment credit. http://www.revenue.nebraska.gov/info/endow_credit.html. Accessed March 11, 2015.

Nebraska Department of Revenue Research Division (2008). State of Nebraska tax expenditure report 2008.

Nichols, M. (2012). U.s. charitable giving approaches $300 billion in 2011. *Reuters*.

Office of Economic Research and Analysis (2014). The revenue impact of Arizona's tax expenditures. Arizona Department of Revenue.

Office of Fiscal Analysis (2012). Connecticut tax expenditure report.

Office of Management and Budget (2017). America first a budget blueprint to make america great again. Technical report, Executive Office of the President of the United States.

Okla. Admin. Code §710:50-15-113. Qualified independent biomedical research institute or qualified cancer research center credit. http://www.oar.state.ok.us/viewhtml/710_50-15-113.htm. Accessed March 11, 2015.

Okten, C. and Weisbrod, B. A. (2000). Determinants of donations in private nonprofit markets. *Journal of Public Economics*, 75(2):255–272.

Ostry, M. J. D., Berg, M. A., and Tsangarides, M. C. G. (2014). *Redistribution, inequality, and growth*. International Monetary Fund.

Oversight Division (2011). Sunset review of the food pantry tax credit program.

Panizza, U. et al. (2001). Convergence in income inequality. *Journal of Income Distribution*, 10(1-2):1–1.

Partridge, J. S., Partridge, M. D., and Rickman, D. S. (1998). State patterns in family income inequality. *Contemporary Economic Policy*, 16(3):277–294.

Partridge, M. D., Rickman, D. S., and Levernier, W. (1996). Trends in us income inequality: evidence from a panel of states. *The Quarterly Review of Economics and Finance*, 36(1):17–37.

Payne, A. A. (1998). Does the government crowd-out private donations? new evidence from a sample of non-profit firms. *Journal of Public Economics*, 69(3):323–345.

Payne, A. A. (2001). Measuring the effect of federal research funding on private donations at research universities: is federal research funding more than a substitute for private donations? *International Tax and Public Finance*, 8(5-6):731–751.

Payne, A. A. and Smith, J. (2015). Does income inequality increase charitable giving? *Canadian Journal of Economics*, 48(2):793–818.

Peloza, J. and Steel, P. (2005). The price elasticities of charitable contributions: a meta-analysis. *Journal of Public Policy & Marketing*, 24(2):260–272.

Piketty, T. (2014). *Capital in the twenty-first century*. The Belknap Press of Harvard Univeristy Press, Cambridge Massachusetts.

Potters, J., Sefton, M., and Vesterlund, L. (2005). After you—endogenous sequencing in voluntary contribution games. *Journal of public Economics*, 89(8):1399–1419.

Quah, D. (1993). Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics*, pages 427–443.

Ramnath, S. (2013). Taxpayers' responses to tax-based incentives for retirement savings: Evidence from the saver's credit notch. *Journal of Public Economics*, 101:77 – 93.

Randolph, W. C. (1995). Dynamic income, progressive taxes, and the timing of charitable contributions. *Journal of Political Economy*, 103(4):709–738.

Ravallion, M. (2003). Inequality convergence. *Economics Letters*, 80(3):351–356.

Revenue Research Division (2013). Biennial tax expenditure report,. North Carolina Department of Revenue.

Sack, J. L. (2001). Proposal would expand americorps for security role. *Education Week*.

Simmons, W. O. and Emanuele, R. (2010). Are volunteers substitute for paid labor in nonprofit organizations? *Journal of Economics and Business*, 62(1):65–77.

Simon, C. A. (2002). Testing for bias in the impact of americorps service on volunteer participants: Evidence of success in achieving a neutrality program objective. *Public Administration Review*, 62(6):670–678.

Simon, C. A. and Wang, C. (2002). The impact of americorps service on volunteer participants results from a 2-year study in four western states. *Administration & Society*, 34(5):522–540.

Solt, F. (2016). The standardized world income inequality database. *Social science quarterly*, 97(5):1267–1281.

Surveillance, Epidemiology, and End Results (SEER) Program. Populations (1969-2012) (www.seer.cancer.gov/popdata). National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released March, 2014.

Tax Analysis Division, Office of Revenue and Tax Analysis (2014). Community/education foundation and homeless shelter/food bank credits 2014 report. http://www.michigan.gov/documents/treasury/ CommunityEducation-FoundationHomlessShelterFoodBankCredits_2014_462174_7.pdf. Accessed March 10, 2015.

The Editorial Board (2014). Broken promises on national service.

The Tax Policy Division of The Oklahoma Tax Commission (2012). Tax expenditure report 2011-2012. http://www.tax.ok.gov/reports/TaxAccessed March 11, 2015.

The Urban Institute. NCCS Core File (Public Charities [2004-2011]).

Tinkelman, D. (2010). Revenue interactions: Crowding out, crowding in, or neither. In Seaman, B. A. and Young, D. R., editors, *Handbook of research on nonprofit economics and management*, chapter 2, pages 18–41. Edward Elgar Publishing.

U.S. House Committee on the Budget (2013). Report of the committee on the budget house of representatives to accompany h. con. res. 25.

Vesterlund, L. (2003). The informational value of sequential fundraising. *Journal of Public Economics*, 87(3):627–657.

West Virginia State Tax Department (2012). West Virginia tax credit review and accountability report - 2012.

Biography

Daniel Teles is a Ph.D. candidate in the Department of Economics at Tulane University, where he has served as a Murphy Graduate Fellow and a Graduate Community-Engaged Fellow within the Center for Public Service. Born in Plymouth, Massachusetts in 1983, he holds a B.S. in Economics from The George Washington University and an M.A. in Economics from Tulane University. Dan started the Ph.D. program at the Tulane University economics department in August 2011. His research focuses on the effects of government policies on the nonprofit sector and the measurement of inequality. He has presented at the Society of Government Economists Annual Conference, the Southern Economic Association Annual Conference, the National Tax Association Annual Conference on Taxation, and at Tulane's Social Justice and the City: City, Culture, and Community Symposium.