Regulatory Hurdles and Growth of Charitable Organizations: Evidence From a Dynamic Bunching Design

Benjamin M. Marx*

November 19, 2012

JOB MARKET PAPER

Abstract

Taxes and regulations, such as labor laws and reporting requirements, often exempt small firms, creating incentives to stay small or delay growth. Firms' responses to such size thresholds provide an opportunity to empirically assess consequences of regulations and firms' willingness to pay to avoid them. This paper presents a new dynamic research design to estimate income responses to thresholds and analyzes an income notch at which IRS reporting requirements for charitable organizations become more onerous. I estimate that the average charity will reduce reported income by $600 to $1000 to "bunch" with those below the notch. In addition, a significant share of charities fail to file when first required to report more information. There is some evidence of retiming of income to delay growing above the notch, but a long-run reduction in the share that grow above the notch provides evidence of real responses as well. Relatively low-expense and low-asset charities are most likely to reduce reported income to stay below the notch, while charities with past receipts above the notch do not manipulate income to get below it, suggesting the report imposes an adjustment cost on new filers. The results highlight the benefits of the dynamic approach, which isolates responses at the time the threshold is encountered by conditioning on past income and growth, for clarifying heterogeneity in responses and estimating extensive-margin responses and long-run effects.

* Department of Economics, Columbia University, 420 West 118th Street, New York, NY 10027. Correspondence should be sent to bmm2126@columbia.edu. I am especially grateful to Wojciech Kopczuk, Bernard Salanié, and Miguel Urquiola for invaluable advice and discussions. I also thank Jonathan Dingel, Jessie Handbury, Todd Kumler, Ilyana Kuziemko, Corinne Low, David Munroe, Giovanni Paci, Petra Persson, Jonah Rockoff, Maya Rossin-Slater, Lesley Turner, Eric Verhoogen, Reed Walker, Caroline Weber, and seminar participants at the Office of the U.S. Department of the Treasury Office of Tax Analysis and National Tax Association Annual Conference. I gratefully acknowledge the financial support of a Nonprofit Research Fellowship from the National Bureau of Economic Research.

1 Introduction

Price discrimination, income eligibility limits, and other policies create "notches" – discontinuities in budget constraints (Slemrod, 2010). Notches create incentives that can distort behavior. A pervasive example is a notch at which expenses rise discretely with income, creating incentives for agents to reduce (reported) income. Such income notches can be found in policies that provide benefits to low-income individuals, such as Medicaid (Yelowitz, 1995), or that restrict government attention to high-income firms, as have elements of the Sarbanes-Oxley Act (Iliev, 2010), the Americans With Disabilities Act (Acemoglu and Angrist, 2001) and Affordable Care Act, and some countries' payroll tax systems (Dixon et al., 2004). Income notches produce deadweight loss if they affect income.

In this paper I show that an Internal Revenue Service reporting notch reduces the incomes of public charities in both the short and the long run. The notch is an income eligibility limit for using simplified IRS reporting forms. The distribution of income exhibits "bunching" of charities at income levels just below the notch. I estimate the number of extra charities below the notch using the techniques of a growing literature that infers behavioral parameters from bunching at kinks and notches. With this approach I estimate that the number of extra charities equals the total number that should be observed with incomes in a range up to $600 above the notch, implying an average willingness to pay $600 to avoid filing the lengthier form. The estimated reduction in the number charities with incomes above the notch, however, is even larger than the excess number below the notch. Because it is difficult to account for attrition or heterogeneity with existing techniques, I develop a new dynamic bunching research design that conditions on past income and income growth. With this new design I affirm the static estimates and show that extensive-margin responses (in this context failure to file the IRS return on time) account for the extra reduction in the number of charities above the notch. In addition, I find that the notch permanently reduces the growth of charities and that smaller charities (in terms of assets or expenses) are the most likely to hold their incomes below the notch. Both ordinary least squares and maximum likelihood estimates show that bunching is confined to charities that were previously below the notch and therefore eligible to file the simplified IRS form, providing evidence that much of the compliance cost imposed by the long form is a one-time adjustment cost to establish the requisite knowledge or financial management infrastructure.

The reporting notch for public charities provides a useful setting for thinking about optimal regulation, particularly as it relates to the large and active charitable sector. In the United States, the charitable sector accounts for 9.2 percent of all wages and salaries (Roeger et al., 2012). Example charities of the size studied in this paper include arts organizations, athletic leagues, economic and social development programs, and youth organizations. The IRS exempts public charities from corporate income taxes in exchange for their commitments to provide social benefits and not distribute profits.[1] Tax exemptions for charities, and tax deductions for donors, create opportunities for tax avoidance and evasion. About a third of each annual IRS "Dirty Dozen" list of tax evasion schemes involves public charities. To enable monitoring by the government and other stakeholders, charities of sufficient size must annually file IRS Form 990, an information return. For most of the last two decades, those with gross receipts below $100,000 and year-end assets below $250,000 were eligible to file the simpler Form 990-EZ, creating a reporting notch in both receipts and assets. The optimal design of such a regulatory notch depends on the extent of social benefits obtained by monitoring charities and costs that monitoring imposes. Recent increases in IRS monitoring of the nonprofit sector and ongoing Congressional hearings demonstrate renewed interest in the optimal regulation of charitable organizations.[2] The optimal design of the monitoring notch for public charities reflects trade-offs that arise in regulation more broadly and is useful for thinking about notches that impose heterogeneous costs.

One contribution of this paper is to extend the methodology of bunching estimation used in a growing body of empirical work. Saez (2010) showed that the extent of bunching (i.e. excess mass) in the distribution of income around a tax schedule kink reveals the tax price elasticity of income. Bunching has since been estimated at kinks in the tax schedule (Chetty et al. (2011)) and at notches in taxes (Kleven and Waseem, 2012) and regulatory schedules (Sallee and Slemrod, 2010).[3] The identifying assumption in bunching estimation is that the distribution of income would be smooth if not for the threshold (whether a notch or kink). The researcher can therefore use observations away from the threshold to construct a counterfactual income distribution. The difference between the mass observed near the threshold and the mass predicted by the counterfactual provides an estimate of the share of agents who bunch.

To supplement the standard approach I develop a dynamic version of bunching estimation. As in static estimation, the goal is to quantify distortions in what would otherwise be a smooth distribution of income. Rather than restricting attention to the univariate distribution of current income, the dynamic approach exploits panel data to identify distortions in the joint distribution of income in multiple periods. To implement the dynamic approach I estimate distributions of growth conditional on current income, comparing charities

[1] The "nondistribution requirement" prohibits nonprofits from paying operating profits to individuals who exercise control over the organization. Excise taxes can be imposed on "excess benefit transactions" including compensation packages deemed to be excessive. Nonprofits include foundations, churches, political groups, and labor organizations in addition to the public charities studied in this paper. State laws vary but frequently exempt charities from income and sales taxes.

[2] Reforms since 2007 include requiring individuals to maintain receipts for noncash donations, revising the 990 forms for fiscal years 2008 and after to require more information from each organization, and in 2011 revoking the tax-exempt status of more than a quarter-million organizations that had not filed in the past three years. In an October 6, 2011 letter to the IRS Commissioner, House of Representatives Committee on Ways and Means Chairman of the Subcommittee on Oversight Charles Boustany wrote that members of both the Oversight and Health Subcommittees "have expressed concern that other tax-exempt organizations may not be complying with the letter or the spirit of the tax-exempt regime, yet continue to enjoy the benefits of tax exemption." In 2012 the Subcommittee Chairman called a series of hearings to elicit testimony from the IRS and experts on the nonprofit sector, and the IRS will be holding a public hearing on proposed regulations or charitable hospitals.

[3] Other recent papers estimating bunching include Bastani and Selin (2012), Carillo et al. (2012), Kopczuk and Munroe (2012), Ramnath (2012), and Weber (2012). Kleven and Waseem (2012) also build on the theoretical work of Saez (2010), deriving formulas relating the taxable income elasticity to bunching at tax notches and kinks when the bunching is limited by optimization frictions.

approaching the notch to charities with similar growth rates but different current income. I first bin the joint distribution to obtain transparent ordinary least squares estimates of income and other responses, testing for heterogeneous responses and for long-run effects on growth. I then use a maximum likelihood approach to precisely estimate the extent of income manipulation and extensive-margin responses.

Dynamic bunching estimation offers several potential benefits. First, conditioning on current income and other variables as in my dynamic approach makes it easier to distinguish which variables predict income responses and which respond concurrently with income. By comparing a treatment group that approaches the notch to control groups with similar growth rates but different initial income, the researcher can identify the effect of approaching the notch on a charity's behavior and test for heterogeneity. Second, it is possible to directly identify long-run effects of a notch on income growth. Third, one can estimate extensive-margin responses or other sources of sample selection related to the threshold. Fourth, identification relies on assumptions that are arguably more plausible than those of the static approach, particularly for responses to a notch that agents face repeatedly. Repeated bunching or persistence of income from year to year could generate growing distortions in the cross-sectional income distribution used in static estimation, but my approach accounts for such dynamics by conditioning on income in the year prior to the year that a charity approaches the notch and identifying distortions in the distribution of growth rates.

A few papers have studied dynamic aspects of bunching. For example, thresholds in time may induce bunching in intertemporal decisions such as the choice of when to claim retirement (Manoli and Weber, 2011), but these one-time decisions will not generate the repeated bunching or long-run effects identified here. Gelber et al. (2012) examine whether bunching persisted after elimination of the Social Security earnings test but not how the policy affected earnings dynamics while it was in place. Kleven and Waseem (2012) use panel data to estimate the share of taxpayers remaining just above or below tax notches. I present a version of such analysis, taking the additional step of estimating a counterfactual probability of remaining in place, as an example of how panel data can be used to explore heterogeneity within the static framework. More similar to my dynamic design is the work of Schivardi and Torrini (2007), who look for distortions in growth rates around a 15-employee notch in Italian labor law. They estimate that the probability of positive growth is reduced by 2 percent for firms in a bin just below the notch, assume these firms would grow to the bin just above the notch, and construct a counterfactual size distribution by solving for the steady-state of an adjusted one-year transition matrix between employment levels. My design follows a similar logic but enables analysis of a continuous variable, estimates bunching and the counterfactual distribution of growth from each level of current receipts, and provides tests for heterogeneity and threshold-related attrition.

The findings of this study provide new information about the behavior of charities and more broadly about the growth of firms and threshold policies. Just as fiscal policy instruments may affect the long-run distribution of wealth, regulatory instruments may affect the firm size distribution and its evolution. The responses of charitable organizations to the IRS filing threshold produce clear distortions in the distribution of reported income. Average income is reduced by several hundred dollars per charity in a neighborhood of the notch, and the share growing income to a point above the notch is significantly reduced for a decade or more. Moreover, policy effects interact with measures of organizational capacity similar to those that have been shown to influence the evolution of the for-profit firm size distribution (Cabral and Mata (2003); Angelini and Generale (2008)). Controlling for current income, a one percent increase in a charity's expenses or assets is associated with a 2.5 percent reduction in the probability of manipulating receipts when approaching the notch in the next year.

This paper also contributes to the literature on firm compliance costs by providing evidence that charities manipulate income to avoid incurring the adjustment cost of complying with new reporting requirements. Tax and regulatory compliance costs made up close to three percent of the revenue of the 1300 largest firms in 1992 (Slemrod and Blumenthal, 1993). Compliance costs appear to have an important fixed component because their burden is proportionately heavier on smaller businesses (Slemrod and Venkatesh, 2002). The estimates in these papers preceded the Sarbanes-Oxley Act, which greatly increased reporting requirements. Public charities also face scale economies in compliance, which consumes 7 percent of the annual budgets of surveyed charities with revenue below $100,000 (Blumenthal and Kalambokidis, 2006). Consistent with these findings, I provide evidence that adjustment is an important component of total compliance cost. I find that charities whose incomes in the prior year necessitated filing a long form showed no propensity to reduce current income by even a small amount to avoid filing again.

The benefit of imposing reporting costs is that firms must disclose information for use by the government and individuals. Investors in for-profit firms appear to value mandatory disclosure of financial information (Greenstone et al., 2006), and the same is likely true of donors to nonprofit firms. The reporting notch therefore reflects a trade-off between imposing additional compliance costs on charities and obtaining additional information from them, much like the calculus of weighing compliance and administrative costs against tax revenues when setting a VAT tax that excludes small firms (Keen and Mintz (2004), Dharmapala et al. (2011)). It is known that income responses must be considered in such situations, and I derive a formula for welfare effects of setting regulatory notches when responses include avoidance and evasion.

The paper proceeds as follows. Section 2 describes the empirical setting of an IRS filing notch for public charities and provides a simple welfare model that motivates estimation of the bunching ratio. In Section 3 I replicate the static approach in the literature and discuss prospects for exploiting panel information within the static design. Section 4 describes the general concept of dynamic bunching estimation. Section 5 presents a reduced-form dynamic approach to estimating responses to notches and testing for heterogeneity in the degree and kind of responses. Section 5 demonstrates a maximum likelihood estimation strategy to precisely estimate the extent of the response and allow for attrition that may be endogenous to the notch. Section 6 concludes.

2 The Setting: Nonprofit Information Returns

This section describes the context of a reporting notch for U.S. charities, the data on these firms, and a simple model for determining the government's optimal notch policy. The importance of the sector, the existence of longitudinal data, and the current interest in regulation of charities make this setting an attractive application for dynamic bunching estimation.

2.1 Background on the Reporting of Charities

The charitable organization reporting notch provides an excellent application for dynamic bunching estimation because: (1) IRS forms create a clear notch in compliance costs, (2) charitable firms face the notch year after year, (3) there is roughly twenty years of longitudinal data on these organizations, and (4) regulatory hurdles will reduce welfare if they discourage these firms from growing to serve their social missions.

I examine IRS reporting notches for public charities. Public charities are organizations granted income and sales tax exemption under section 501(c)(3) of the Internal Revenue Code. All public charities with gross receipts over \$25,000 (except religious congregations) must annually file information returns with the IRS using Form 990 or Form 990-EZ. For fiscal years starting before 2008, charities with gross receipts exceeding \$100,000 or year-end total assets above \$250,000 were required to file the lengthier Form 990.[4]

Form 990 requires charities to access and report more financial data than Form 990-EZ. Table 1 presents a comparison of 990-EZ and 990 for fiscal years beginning in 2007 or earlier. The two forms require nearly all the same categories of information, but Form 990 requires much greater detail. Form 990 contains more lines in most sections and requires a detailed statement of functional expenses. Estimates under the Paperwork Reduction Act for the time required for completion and filing are 164 hours for Form 990-EZ and 260 hours for Form 990 (Internal Revenue Service, 2007). The time estimates include the required Schedules A and B and include time required to perform the necessary recordkeeping (the majority of the difference between the two forms), to learn about the forms, and to prepare and assemble them. The raw difference of roughly 100 hours (a 59% increase), if accurate for the marginal charity near the notch, would imply that an organization

[4]The IRS also provides simplified individual income tax forms for filers with incomes below a notch, but it turns out this notch is not sufficiently relevant to observe bunching in the distribution of individual incomes. While eligibility for filing Form 1040-EZ is restricted to taxable incomes below \$100,000, other restrictions on age, types of income, and filing status restrict its use among filers even if their incomes are below the notch. Inspection of the distribution of incomes among filers in the IRS Tax Model data reveals very few 1040-EZ filers with income near the notch.

with receipts above $100,000 by less than 100 times the hourly wage could forgo enough receipts to stay below the notch and have more net resources as a result. Blumenthal and Kalambokidis (2006) asked for the titles and qualifications of individuals responsible for filings and imputed hourly wages between $13.09 and $51.77. If all charities faced a marginal cost of filing Form 990 equal to 100 hours at a rate of $13 per hour then none should report receipts between $100,000 and $101,300. Realistically, the marginal cost of filing would vary with the amount of recordkeeping already being performed, implying variation in the amount of receipts charities would forgo to avoid filing. Blumenthal and Kalambokidis (2006) also find that after controlling for size and other factors, those filing Form 990 report spending about 45% more on professional advisory fees than those filing Form 990-EZ.

Form 990 may also impose a disclosure cost on charities that do not want to reveal certain information. For example, charities filing Form 990 must check a box if any officers or key employees are related to each other and must list any former officers that were compensated during the year. However, most potentially-sensitive information is required of both types of filers: compensation of current officers and employees must be listed on each form, and the rule for completing Schedule B (Schedule of Contributors) is the same for both forms.[5] Moreover, charities near the eligibility notch at $100,000 of gross income are unlikely to be able to pay large salaries. It will not be possible to fully test for disclosure costs, but I look for suggestive evidence by relating income manipulation to ex-post values of items appearing only on Form 990.

Income threshold policies may create incentives for entities to reorganize as multiple smaller organizations (Onji, 2009). In the present context this incentive is likely to be weak because exempt status would have to be applied for and obtained for each organization and because economies of scale are likely to be considerable at sizes small enough to make organizations eligible to file Form 990-EZ. I therefore treat each charity as an individual unit.

2.2 Panel Data on Charities

This study uses IRS data from the Core files of the National Center for Charitable Statistics (NCCS), a division of the Urban Institute. IRS databases offer the most comprehensive standardized data on tax-exempt organizations in the U.S. The IRS produces a Business Master File of descriptive information from each filing and Return Transaction Files of financial information. The NCCS Core files contain data from the IRS databases on all 501(c)(3) organizations that were required to file a Form 990 or Form 990-EZ and

[5] Public charities must file Schedule B if they received any individual contributions of more than $5000. Those meeting the "public support test" of receiving more than a third of their support from general, public sources must also file Schedule B if they receive an individual contribution greater than 2% of total contributions. Amounts and descriptions from a public charity's Schedule B are made available for public inspection, but information identifying contributors is not.

complied in a timely manner.[6] Analyses "show the IRS 990 Returns to be a generally reliable source of financial data," although inattention by filers adds noise to the data and purposeful expense shifting may inflate program-related expenses relative to administrative expenses (Froelich and Knoepfle, 1996). This study makes limited use of expense categories and explicitly examines manipulation of revenue around the Form 990 reporting notch.

Several financial variables from each form appear in the data. In this paper I focus on gross receipts. "Contributions, gifts, grants, and similar amounts received" make up the largest component of gross receipts. The other components are program service revenue, membership dues, investment income, gross sales of inventory, gross sales of other assets, and other revenue, all of which appear in the data. Total assets, liabilities, and expenses are each available for both types of filing. While both forms require listing all officers, directors, and trustees and the compensation paid to each, compensation only appears in the data for organizations that filed Form 990.[7] Other variables populated for all filings include the date at which tax-exempt status was granted, reasons for 501(c)(3) status, and codes describing the type of organization and services provided. I do not use the limited set of variables collected from Schedule A, which includes lobbying and other political expenses that equal zero for a large majority of organizations.

I analyze public charities in filing years 1990 to 2010, the years for which data on public charities are currently available. Marx (2012) compiled data on private charitable foundations going back to the 1960s, but private foundations file Form 990-PF and hence do not face the same notch as public charities. Data for each NCCS file year comprise the most recent return filed by each organization. Unfortunately, the variable indicating whether organizations filed Form 990 or 990-EZ is not available for file years preceding 2006. I use the Form 990 variable to show that the receipts notch is a binding constraint for many charities in 2007 but use observations from the earlier years throughout the analysis.

Table 2 provides summary statistics showing the prevalence of small charities. The $100,000 receipts notch (which has been defined nominally and not adjusted for inflation) falls between the lower quartile and median of gross receipts. Expenses are highly correlated with gross receipts, while assets exhibit greater variation. Of the more than four million observations in the data, over 20,000 have receipts in a region around the notch. The IRS and NCCS classify charities according to the National Taxonomy of Exempt Entities, which groups charities into major and minor categories. Education is the most common major category among organizations near the notch, of which many fall into minor categories indicating organizations that support schools. Other charities of this size include religious groups, arts organizations, and athletic leagues.

[6]To create the Core file, NCCS cross-checks and cleans data from the various IRS databases and from organizations' 990s when necessary. NCCS carries out a variety of procedures to check and clean the data. A detailed description of the Core Files and other data is available from the National Center for Charitable Statistics (2006).

[7]Form 990 contains separate lines for compensation of current officers and directors, former officers and directors, and other employees, while Form 990-EZ contains just one line for "Salaries, other compensation, and employee benefits."

Figures 1 and 2 show that the filing notch binds, and charities bunch below the notch. Figure 1 shows that, for charities with fiscal years that begin in 2007, the probability of filing Form 990 is discontinuous at the receipts notch. Just under half of organizations with receipts just below the $100,000 notch file Form 990-EZ. About 17% of firms in this region must file the longer form because their assets are above the $250,000 notch. The others to file Form 990 by choice, perhaps to satisfy donors or because they had filed it in the past. The fact that some firms choose to voluntarily file Form 990 suggests heterogeneity in organizations' cost structures or preferences. Since recordkeeping accounts for much of the estimated cost difference between the two types of filing, organizations that have already made the necessary investment in their administrative capacity would find it less costly to switch to the longer form. Among those with 2007 current receipts below the notch and 2006 receipts above, nearly 80 percent continue to file Form 990. In the empirical analysis I present further evidence that adjustment is a primary component of the compliance cost, with organizations that have previously filed the long form showing little propensity to bunch below the notch. The fact that a considerable share of organizations file Form 990 before reaching the notch should be kept in mind when interpreting results but does not affect the analysis except for the fact that it will not be possible to identify a strictly dominated income region as in the work of Kleven and Waseem (2012).[8]

This study analyzes income responses to the notch.[9] Figure 2 shows a histogram of receipts. The distribution of receipts is smooth except for an excess of mass just below the notch. This excess of mass of bunchers is the object of interest, as supported by the model in the next section. Charities must also file Form 990 if their assets exceed $250,000, but bunching at this asset notch is less conspicuous. Tests suggest a small discontinuity in the density of assets with statistical significance that is sensitive to the choice of bin width. The asset notch is binding for fewer organizations, since roughly 72% of charities with assets between $200,000 and $250,000 have receipts over $100,000, and an additional 15% in this range file the full Form 990 by choice. I therefore focus on the receipts notch in the model and empirical analysis.

2.3 A Simple Model of Welfare Effects of a Notch Policy

This section presents a welfare analysis to provide a conceptual framework for evaluating bunching estimates. The model demonstrates that the optimal location of the notch depends on the counterfactual density near the notch and the excess bunching mass below it, quantities estimated in other bunching studies to measure the taxable income elasticity. In the model, policy design weighs the social value of obtaining information

[8]Charities filing Form 990 by choice are relatively young, rapidly growing, high-expense, and with most receipts in the category "Contributions, gifts, grants, and similar amounts received." Variation across states shows no clear pattern; Illinois, Maine, and Pennsylvania have auditing requirements that apply to charities with contributions at levels below $100,000 but do not have a significantly higher share of Form 990 filers.

[9]If the level of receipts was exogenously determined then Figure 1 could represent the first stage in a fuzzy regression discontinuity study of the effect of Form 990 on, say, donations received. Since organizations can manipulate their receipts by varying fundraising expenditures or shifting receipts across years, regression discontinuity is not appropriate.

through reporting requirements against the costs these requirements impose, including the avoidance costs of organizations that bunch.

The Charity

A charity seeks to maximize expendable net income x. The charity earns receipts (income) y and reports receipts $r := y - a$ to the government, where the amount a is kept hidden by tax avoidance or evasion. The total cost to the charity is the sum of the cost $A(y, a, \gamma, \omega)$ of avoidance and the cost $B(y, \omega)$ of earning the amount y in receipts, where γ is a vector of parameters describing heterogeneity and ω is a parameter describing heterogeneity in fundraising ability that is unrelated to (but perhaps correlated with) the cost of avoidance. Assume the cost functions are nondecreasing and convex in y and a and that $A_y(y, 0, \gamma, \omega) = 0$. This formulation is in keeping with the "general model of behavioral response to taxation" of Slemrod (2001); the cost of avoidance includes both direct psychic or financial costs as well as changes to the expected cost of an audit, and avoidance opportunities may vary with income. The organization must also pay filing cost $\phi(\gamma, \omega)$ if $r > \rho$, the filing threshold. The budget constraint is thus $x \leq y - A(y, a, \gamma, \omega) - B(y, \omega) - \phi(\gamma, \omega) \cdot 1\{r > \rho\}$, and the firm's problem is

$$\max_{y,a} \{y - A(y, a, \gamma, \omega) - B(y, \omega) - \phi(\gamma, \omega) \cdot 1\{y - a > \rho\}\}$$

If the filing constraint does not bind then optimal avoidance is zero, and the first-order condition $B_y(y, \omega) = 1$ defines the optimal value of receipts $\bar{y}(\omega)$ as that level of fundraising at which the marginal cost of raising one dollar has reached one dollar. Because $\bar{y}(\omega)$ plays an important role throughout the analysis, from this point I simply describe fundraising heterogeneity in terms of $\bar{y}$. There will be a one-to-one relationship between $\bar{y}$ and ω if $\frac{d\bar{y}}{d\omega} = -\frac{B_{y\omega}(y,\omega)}{B_{yy}(y,\omega)} > 0$, implying that the inverse function $\omega(\bar{y})$. I therefore rewrite $\phi(\gamma, \omega)$ as $\phi(\gamma, \bar{y})$ and define $C(y, a, \gamma, \bar{y}) = A(y, a, \gamma, \omega(\bar{y})) + B(y, \omega(\bar{y}))$.

If the filing constraint does bind, i.e. optimal reported income is $r = \rho$, then $y - a = \rho$, and the problem becomes

$$\max_{y} \{y - C(y, y - \rho, \gamma, \bar{y})\}$$

In this case the first-order condition gives $C_y(y, a, \gamma, \bar{y}) = 1 - C_a(y, a, \gamma, \bar{y})$. Receipts fall short of $\bar{y}$ because marginal earnings increase the necessary amount (and therefore cost) of avoidance. Call the level of receipts that satisfies this condition $\hat{y}(\gamma, \bar{y})$, which I will generally write simply as $\hat{y}$.

When will the charity bunch at the reporting threshold? If $\bar{y} \leq \rho$ there is no need to misreport. If $\bar{y} > \rho$ then the charity obtains $\hat{y} - C(\hat{y}, \hat{y} - \rho, \gamma, \bar{y})$ if it reports $r = \rho$ and $\bar{y} - C(\bar{y}, 0, \gamma, \bar{y}) - \phi(\gamma, \bar{y})$ if it does not. The charity will therefore bunch if and only if $\phi(\gamma,\bar{y}) \geq (\bar{y}-\hat{y}) - [C(\bar{y},0,\gamma,\bar{y}) - C(\hat{y},\hat{y}-\rho,\gamma,\bar{y})]$. Because costs are convex we can define $\delta(\rho,\phi,\gamma,\bar{y})$ as the maximum difference (possibly zero) between $\bar{y}$ and ρ from which the organization would be willing to choose $r=\rho$. That is, a charity bunches at the notch if $\rho < \bar{y} \leq \rho + \delta(\rho,\phi,\gamma,\bar{y})$. Again simplifying notation, I will suppress the arguments of δ.

Reported receipts are

$$r = \begin{cases} \bar{y} & \bar{y} \leq \rho \\ \rho & \rho < \bar{y} \leq \rho + \delta \\ \bar{y} & \bar{y} > \rho + \delta \end{cases}$$

The charity obtains indirect utility

$$V(\rho,\phi,\gamma,\bar{y}) = \begin{cases} \bar{y} - C(\bar{y},0,\gamma,\bar{y}) & \bar{y} \leq \rho \\ \hat{y} - C(\hat{y},\hat{y}-\rho,\gamma,\bar{y}) & \rho < \bar{y} \leq \rho + \delta \\ \bar{y} - C(\bar{y},0,\gamma,\bar{y}) - \phi & \bar{y} > \rho + \delta \end{cases}$$

Note that ρ enters directly for bunchers but not others. This implies that changes to the location of the threshold will have first-order effects on the utility of inframarginal bunchers (but not others).

The Government

The government's problem is to maximize the net value of the reporting regime. Social welfare includes the indirect utility of charities as well as the (external) social benefit obtained from reporting. The social benefit of an organization's disclosure spending, net of the administrative cost to the government, is $\pi(\phi,\gamma,\bar{y})$. Potential income is distributed with cumulative distribution function $F(\bar{y})$ and probability density function (pdf) $f(\bar{y})$. The heterogeneity parameter γ has pdf $g(\gamma)$. Social welfare per firm[10] is

$$W = \int \left[\int_0^\infty V(\rho,\phi,\gamma,\bar{y}) f(\bar{y})\, d\bar{y} + \int_{\rho+\delta}^\infty \pi(\phi,\gamma,\bar{y}) f(\bar{y})\, d\bar{y} \right] g(\gamma)\, d\gamma$$

$$= \int \Bigg[\int_0^\infty (\bar{y} - C(\bar{y},0,\gamma,\bar{y})) f(\bar{y})\, d\bar{y} + \int_{\rho+\delta}^\infty (\pi(\phi,\gamma,\bar{y})) f(\bar{y})\, d\bar{y}$$

$$- \int_{\rho+\delta}^\infty \phi(\gamma,\bar{y}) f(\bar{y})\, d\bar{y} + \int_\rho^{\rho+\delta} [\hat{y} - \bar{y} - (C(\hat{y},\hat{y}-\rho,\gamma,\bar{y}) - C(\bar{y},0,\gamma,\bar{y}))] f(\bar{y})\, d\bar{y} \Bigg] g(\gamma)\, d\gamma$$

[10]Donor utility is excluded from the social welfare function, as recommended in research on optimal taxation of charitable giving (e.g., Andreoni (2006), Diamond (2006)). In addition to their arguments there is evidence that fundraising reduces the utility of the average prospect (DellaVigna et al., 2012).

With the social welfare function written as the sum of these four terms, one can immediately see how policy will affect social welfare. Policy-makers can influence two parameters, the location of the notch and the cost of reporting. Increasing the amount or complexity of information reported on the long form will increase ϕ. From terms two and three one sees that this will directly increase welfare to the extent that this new information is of net social benefit but will reduce the number of number of charities filing the long form. The choice of how much detail to require in financial reports is therefore similar to optimal screening of social benefits under imperfect takeup because greater complexity has direct benefits but may reduce participation (Kleven and Kopczuk, 2011). I will not attempt to estimate the social value of reporting.[11] It turns out, however, that the optimal location of the threshold depends on estimable quantities analogous to those studied in the tax bunching literature. Marginal changes to ρ will affect all but the first term in the social welfare function, but marginal bunchers (with $\bar{y} = \rho + \delta$) experience no first-order utility changes due to the indifference condition $\phi(\gamma, \bar{y}) = ((\rho + \delta) - \hat{y}(\rho + \delta)) - [C(\rho + \delta, 0, \gamma, \bar{y}) - C(\hat{y}(\rho + \delta), \hat{y}(\rho + \delta) - \rho, \gamma, \bar{y})]$ and indifference for those with $\bar{y} = \rho$. After using the indifference conditions to cancel terms,

$$\frac{dW}{d\rho} = \int \left[\int_{\rho}^{\rho+\delta} C_a(\hat{y}(\bar{y}), \hat{y}(\bar{y}) - \rho, \gamma, \bar{y}) f(\bar{y})\, d\bar{y} \right] g(\gamma)\, d\gamma - \int (1 + \delta_\rho)\, \pi(\rho + \delta, \phi, \gamma, \bar{y}) f(\rho + \delta) g(\gamma)\, d\gamma$$

Raising the threshold has two counteracting effects. First, charities that were bunching achieve some savings because they no longer have to avoid reporting as much income. Second, raising the threshold reduces the amount of information available to the extent that previously-indifferent charities now bunch at the threshold.

The expression for the welfare effect of moving the notch becomes simpler when written it terms of averages in the region from which bunching occurs. The main identifying assumption in bunching estimation is that bunching is local and there exists some $\bar{\delta} = \max(\delta(\rho, \phi, \gamma, \bar{y}))$. The localness assumption restricts the degree of heterogeneity and would hold, for example, if there is an $M > 0$ such that for all γ, $\bar{y}$ and y we have $C_a(y, 0, \gamma, \bar{y}) \geq M$. Denote the excess mass observed below the notch as $B := \int \left[\int_{\rho}^{\rho+\delta} f(\bar{y})\, d\bar{y} \right] g(\gamma)\, d\gamma = b\left(F\left(\rho + \bar{\delta}\right) - F(\rho)\right)$, where b is the share of organizations that choose to bunch. Assume that $\frac{db}{d\rho} \approx 0$, $\frac{d\bar{\delta}}{d\rho} \approx 0$, and $\exists \bar{\pi} : \forall \gamma, \bar{y} \in [\rho, \rho + \delta], \pi(\phi, \gamma, \rho + \delta) \approx \bar{\pi}$. In words, slight movements of the notch have little effect on the share of organizations that bunch, the maximum amount by which they will reduce income, or the social value of the average buncher's report. The first two assumptions are effectively the same as the simplifications common in the taxable income bunching literature, while the third is useful here due to the

[11] Potential benefits would include reductions in avoidance/evasion on other margins. Examinations of tax-exempt organizations in 1998 through 2005 resulted in recommended additional tax payments (for taxable transactions including payroll and unrelated business income) averaging $106 million per year (Internal Revenue Service, 1998-2005). I have found no sources that present enforcement statistics by form filed.

potential heterogeneity in the social value of reporting.

The simplifying assumptions make it possible to rewrite the term describing the welfare effect of lost reports as $-\bar{\pi}\left[bf\left(\rho+\bar{\delta}\right)+(1-b)\,f(\rho)\right]$.[12] That is, the value of long forms lost is the product of their average value and the change in the share of charities that bunch. Under the simplifying assumption that bunching is proportional to mass in the reduced range, the change in the share of charities that bunch is the weighted average of the values taken by the underlying density at the top and bottom of the reduced range. The welfare criterion for the optimal location of the notch is thus

$$\frac{dW}{d\rho} \geq 0 \Leftrightarrow \frac{\bar{\pi}}{E\left[C_a\left(\hat{y},\hat{y}-\rho,\gamma,\bar{y}\right)|\rho<\bar{y}\leq\rho+\delta\right]} \leq \frac{\left(F\left(\rho+\bar{\delta}\right)-F(\rho)\right)}{\left(bf\left(\rho+\bar{\delta}\right)+(1-b)\,f(\rho)\right)}$$

The expression for the welfare effects of moving a regulatory notch includes factors comparable to those arising from the choices of marginal income tax and VAT rates studied in the literature. When administrative costs increase with the number of covered firms, the optimal income exemption threshold for a value-added tax will induce bunching if the revenue effects are small (Dharmapala et al., 2011). The net benefit of reporting $\bar{\pi}$ plays a role similar to that of tax revenue, although this benefit varies across organizations (as reflected in the fact that only larger organizations are required to file the long form). The regulatory problem is similar to setting a minimum wage or tax rates at low incomes, where extensive-margin responses are likely more important than intensive-margin responses (Saez, 2002). The expression on the right-hand side of the inequality is a version of the ratio that arises in other bunching studies that are motivated by the problem of setting marginal tax rates. The existing literature uses the relationship $b\left(F\left(\rho+\bar{\delta}\right)-F(\rho)\right)=B\approx b\bar{\delta}f(\rho)$ to back out an estimate of $\bar{\delta}$ from estimates of the counterfactual distribution and excess mass. Kleven and Waseem (2012) estimate a parameter similar to b by using the known amount of a tax to identify a strictly dominated region just above the notch, taking those that remain in this region as the share that cannot bunch. I do not observe the exact reporting costs, which I expect to exhibit heterogeneity, and will instead use dynamic techniques to estimate b. Because b is small, and in keeping with other bunching studies, I will report the bunching ratio as the ratio of excess mass to the value of the counterfactual distribution at the notch (rather than the weighted average).

Naturally, it is far more difficult to estimate the marginal cost $C_a\left(\hat{y},\hat{y}-\rho,\gamma,\bar{y}\right)$. The distribution of reported income reveals income responses, but the cost of these responses is not identified without another source of variation. Though the marginal benefits of real and avoidance responses are equated (per the first-order condition) and have the same implications in the model (as is generally true unless externalities

[12] $-\int\left[(1+\delta_\rho)\,\pi(\phi,\gamma,\rho+\delta)\,f(\rho+\delta)\right]g(\gamma)\,d\gamma\approx -\bar{\pi}\int\left[(1+\delta_\rho)\,f(\rho+\delta)\right]g(\gamma)\,d\gamma= \bar{\pi}\frac{d}{d\rho}\int\left[\int_{\rho+\delta}^{\infty}f(\bar{y})\,d\bar{y}\right]g(\gamma)\,d\gamma= \bar{\pi}\frac{d}{d\rho}\left[1-b\left(F\left(\rho+\bar{\delta}\right)-F(\rho)\right)-F(\rho)\right]\approx\bar{\pi}\left[-b\left(f\left(\rho+\bar{\delta}\right)-f(\rho)\right)-f(\rho)\right]=-\bar{\pi}\left[bf\left(\rho+\bar{\delta}\right)+(1-b)\,f(\rho)\right]$.

or other considerations are incorporated), evidence of avoidance is useful in at least two respects. First, relative to a world in which avoidance was prohibitively costly, evidence of avoidance would indicate a lower total cost of manipulating income to stay below the notch, making it less desirable to raise the notch to a higher level of receipts. Second, the extent of avoidance affects inference of the preferences of charities. The amount by which a charity reduces reported receipts in order to bunch provides an upper bound on willingness to pay to avoid reporting because avoidance allows the organization to pay less than the full amount of this reported income reduction. The bound approaches the true value of willingness to pay as the marginal cost of avoidance approaches one. I obtain some evidence of the extent of avoidance by comparing short-run and long-run responses.

The theory in this section motivates the estimation of the excess mass B, the bunching share b, and the counterfactual density at the notch $f(\rho)$. I now turn to estimation of these parameters using static and then dynamic techniques.

3 Benchmark Static Techniques For Estimating Bunching

Before presenting the dynamic bunching research design I follow the static approach used in the literature. I describe the technique for the unfamiliar reader, display the results for public charities, then explore what insights can be obtained within this framework by incorporating other variables, including those requiring panel data.

3.1 The Static Methodology for Estimating Bunching

Bunching empirics exploit distortions in distributions around thresholds at which income or prices change discretely. By estimating the excess mass around a threshold one can obtain reduced-form estimates of policy-relevant behavioral elasticities. Saez (2010) introduced this insight by showing how kinks in marginal tax rates produce a pattern of bunching in the income distribution that reveals the taxable income elasticity without the need to specify a particular utility function. Individuals with incomes above a kink that raises the marginal tax rate have an incentive to reduce reported income, and the greater the income elasticity the more bunching will be observed in the distribution around the kink. Bunching estimation, both at kinks and at notches, quantifies the extent of bunching by comparing the observed distribution to an estimate of the smooth counterfactual that would be expected in the absence of the threshold.

The key to bunching estimation is to construct the counterfactual distribution of income. Static bunching estimates use parts of the density above and below a threshold to construct a counterfactual for the amount of mass that should be at the threshold. Figure 3 provides an example of the static procedure as applied to the

Form 990 filing notch. Most studies approach the distribution as a histogram, constructing bins and plotting the count of observations in each bin as depicted by the circles in Figure 3. The number of observations within some number of bins of the threshold is compared to a counterfactual constructed using bins further away from the threshold. That is, the researcher estimates the counterfactual density by omitting a certain number of bins around the threshold (the "omitted region") and then estimating a smooth function through the values of the other bins. Figure 3 displays the estimated counterfactual as the smooth curve through the data. Some authors construct the counterfactual by taking a simple average of the bin just above the omitted region and the bin just below it, an appropriate counterfactual under the assumption of local linearity of the distribution. Others use a wider range of the distribution and fit polynomials to the bin counts in this range (except for the omitted region around the threshold). In Figure 3 I provide an example using charities with receipts of $50-200,000 and a polynomial of degree 3, which minimizes the Akaike information criterion.

Kinks and notches offer slightly different implications. In the case of a kink, incomes may bunch on either side of the threshold, so the bunching estimate is the sum of all excess mass observed in the omitted region. In the case of a notch, observations that should be on one side of the threshold will instead bunch on the other. For a notch at which costs increase there will be excess mass in the bunching range below the notch and reduced mass in the reduced range above the notch.[13] Both the excess mass and reduced mass can be estimated by comparing the observed density to the counterfactual. Estimation of bunching at kinks requires the further step of raising or lowering the counterfactual distribution on the side of the kink where prices are affected, but this step is not necessary for a notch at which fixed costs are imposed because other than income effects there should be no responses far from the threshold.

3.2 Static Estimation Results

Using the static approach, I estimate significant bunching of public charities at the reporting notch. In the pooled sample a significant excess share of charities appears below the notch and the reduction in the share above the notch is even larger. Annual results show the response of charities to removal of the notch.

Static bunching estimates for the pooled sample appear in Table 3. I use the sample of charities in years up to 2007 that also appear in the prior year (for maximum comparability with the dynamic estimates that follow). The first row of the table shows estimates of excess mass below the notch, the numerator B of the bunching ratio. An estimate of .1 would indicate .1 percent of all charities in that year's sample are below the notch and should be above it. The results from the basic specification, a cubic counterfactual as depicted in Figure 3, indicate that the share of charities appearing below the notch is .148 percentage points greater

[13] I use "reduced" in place of the term "missing" that appears in the literature to maintain a distinction between organizations that shift receipts and those that go missing from the data.

than predicted by the counterfactual. In the second row this number is divided by the value of the density at the notch to give the bunching ratio that is derived from theory in this and other settings. The bunching ratio reveals that the number of bunching charities is roughly equal to the number of charities that should be above the notch by up to $600 (=$100,000*.00592). If all income responses are real, this estimate would imply the average charity is willing to pay $600 to file Form 990-EZ instead of Form 990. The third row displays estimated reduction of mass above the notch, which is .25 percent of the sample.

The estimates in Table 3 raise the question of why the reduction in the number of charities above the notch is significantly larger than the addition below the notch. The basic specification suggests the excess is only about 60% of the reduction, and the size of the reduction suggests charities may be willing to pay as much as $1000 to avoid Form 990. The difference between the estimated excess and reduction could arise due to attrition because, say, Form 990 is more difficult to complete on time and hence more data are missing above the notch. It is not clear how we would test the static specification or whether we should require the excess to equal the reduction. The second and third columns present the results of more flexible specifications that do not reconcile the two results. Allowing for a discontinuity at the notch reduces the estimate on both sides by a very small amount, leaving the asymmetry in the estimates. Estimating a separate polynomial on each side of the notch gives a similar point estimate of the excess but a much larger standard error, and the lack of curvature in the distribution above the reduced range yields an insignificant estimate of the reduced mass. With the dynamic design I will provide support for the basic specification of the static estimate. I will use the distribution of receipt growth to show that the reduction above is greater than the addition below because charities that should be crossing the notch instead drop out of the sample.

Annual static estimates raise another potential question. Figure 4 displays annual cross-sectional estimates of the excess bunching mass.[14] Point estimates are generally around .15, the estimate for the pooled sample, with some variation from year to year. Surprisingly, excess mass remained at incomes just below $100,000 even after the notch was moved to $1,000,000 in 2008. The estimate of .150 for 2008 (with standard error .039) is among the upper half of all the static estimates and significantly different from zero. Gelber et al. (2012) find similar delays in the adjustment of some individuals' earnings to removal of the Social Security earnings test. One potential explanation would be that the extra charities remaining in the bunching region were former bunchers who were unable to raise their receipts rapidly when the notch was moved. An alternative possibility is that the staff of charitable organizations did not understand that the notch had moved and that they continued to actively keep income below $100,000.[15] The dynamic approach offers a way to

[14]For the presented annual results I reduce the allowed bunching range to just $90-100,000, which yields more-tightly-grouped estimates with smaller standard errors.

[15]The revised form for fiscal years starting in 2008 was released in June of 2007, and 99.9 percent of organizations in the NCCS data filed the correct 2008 form for their 2008 fiscal year. While the income eligibility level for the 990-EZ is stated near the top of the form, it is possible that uninformed filers did not look closely and misread the $1,000,000 as the usual $100,000.

distinguish between these possibilities by looking for distortions in the 2008 receipts of charities that were below the notch in 2007. Estimates obtained from the dynamic methodology described in Section 6 support the notion that charities continued to manipulate income in 2008. According to the dynamic approach, an estimated .088 percent of charities (with standard error .041) had receipts below $100,000 in 2007 and chose the remain below $100,000 when they could have crossed. An additional .098 percent (with standard error .035) would have moved to the region just above the notch but instead left the sample.

3.3 Exploring Heterogeneity in the Static Framework

Other variables can provide additional information about heterogeneity in the income response. I show that the static bunching design can exploit panel data to determine whether total bunching is widespread or attributable to a small number of repeated bunchers. I then present growth rates by level of receipts as an example of correlation between income bunching and another variable. These examples indicate some difficulties in interpreting correlations when agents face the same notch repeatedly, which motivates the use of dynamic bunching estimation.

Testing for repeated bunching is straightforward. Repeated bunching is the act of remaining below the notch for more years than expected, so the goal is to estimate the share of charities that would remain near current receipts if the notch did not exist. To estimate repeated bunching I construct bins of current receipts and estimate the probability that in h years the organization remains in its current bin. That is, I partition receipts into bins of width bw and estimate

$$(1) \qquad D\left(bin_i\right)_{i,t+h} = \beta \cdot bunchbin + \sum_{k=1}^{K} \alpha_k r_{it}^k + \gamma_t$$

where $D\left(bin_i\right)_{it+h} = 1\left\{r_{i,t+h} \in bin_i \cap r_{i,t} \in bin_i\right\}$ is an indicator for remaining in the same bin h years in the future, $bunchbin = 1\left\{r_{it} \in [notch - bw, notch)\right\}$ is an indicator for having current receipts in the bunching range, $\sum_{k=1}^{K} \alpha_k r_{it}^k$ is a polynomial in receipts that provides the counterfactual for the bunching range, and γ_t is a vector of year dummies.[16]

I find that charities bunch at the reporting notch for many years. Figure 5 shows the probability of remaining within a $5000 receipts bin three years into the future. This probability varies smoothly with receipts except just below the notch. Among observations in the bin just below the notch, about 5.7 percent remain in the same bin, compared to a counterfactual prediction of only 5 percent.

[16]Since crossing the notch requires a positive growth rate, one could alternatively nonparametrically regress the probability of positive growth on current receipts and estimate any discontinuity at the notch. Such an approach might work well if agents are able to bunch precisely at the exact value of the notch but would underestimate bunching if manipulation is imprecise and bunchers' receipts move around within the bunching range. Manipulation indeed appears to be imprecise around the Form 990 notch, and the choice of bin width should reflect the range of incomes that appear to exhibit bunching. In this setting, similar results obtain for different bin widths and Probit specifications.

Table 4 reports the results from estimating regression (1) for horizons up to 10 years. Observations in the bin just below the notch are about 1.55 percentage points more likely to remain there the following year than would be predicted by surrounding observations. This excess probability of staying in the same range of income declines over time but remains significantly positive for at least five years. Bunching is persistent, suggesting the proclivity to bunch is much stronger in some organizations (the repeat bunchers) than others.

In principle, heterogeneity can be described in the static setting by plotting any other variable as a function of income. Figure 6 presents another example using panel data. The outcome is the growth of log receipts from the previous year to the current year. The average growth rate is discontinuous in receipts, with charities just below the notch having significantly higher growth rates than charities just above. Graphs of this sort may offer a clear interpretation in some cases, but caution is warranted. Here there is an issue of simultaneity: bunching is itself a manipulation of the growth rate, but the growth rate may affect the propensity to bunch. At the same time, plotting variables that are distorted against current income may not reveal the distortion in these variables. Income, for example, would obviously show no discontinuity in itself. Similarly, if expenses are always proportional to reported income then these will also appear undistorted despite responding to the notch in proportion to the income response. In general it will be difficult to tell cause from effect or to disentangle simultaneous responses in multiple variables without making strong assumptions about functional forms.

The static approach provides clear evidence of income responses. Exploiting panel data within the static framework we can also see that observations stay in the bunching region for many years and that growth is distorted at the notch, providing immediate evidence that bunching is a dynamic process in which past income is relevant. I now model the dynamic bunching process explicitly.

4 Introduction to Dynamic Bunching Estimation

Here I describe a dynamic approach to bunching analysis. The idea is to look for distortions not just in the univariate distribution of current income but also in the joint distribution of income in multiple periods. My implementation of dynamic bunching estimation tests for manipulation in the joint distribution of current and future income. Income manipulation is identified under smoothness assumptions about the distribution of growth conditional on current receipts.

When agents face notches repeatedly their choices over time may reveal more information than is captured in the cross section. Because notches and kinks are often fixed in real or nominal terms, they may affect the same agents year after year. These agents therefore face a notch both in current income and future income. With panel data it is possible to observe agents' choices in multiple years and estimate the joint distribution of income over these years.

The intuition of static bunching estimation extends naturally to the multivariate distribution of current and future incomes. For each year in the data, some agents whose incomes should be just above the notch will instead be observed just below the notch. Thus, in the joint distribution of this year's income and next year's income, for example, we should therefore observe bunching of both current incomes and future incomes below the notch and perhaps some interaction between the two periods. To estimate these distortions one must again construct a counterfactual distribution in a neighborhood of the now-multi-dimensional notch.

Relative to the univariate distribution of current income, the joint distribution of current and future income offers the researcher more options. A straightforward option appearing in the literature is to ignore intertemporal correlations and simply pool years as repeated cross sections and perform static estimation. More generally, one could simply estimate the entire multivariate density of income in each year. Allowing such generality, however, would be computationally expensive. To simplify the analysis while retaining potentially valuable correlations the researcher could transform, segment, or collapse the distribution. For example, the univariate procedure could be applied to the subsequent income of agents with current incomes in a range of interest. Regardless of the implementation choice, the the empirical strategy will seek to estimate distortions around the notch in the relevant dimensions.

I perform dynamic estimation in terms of the distribution of log current receipts and growth to the next year's log current receipts. The joint distribution of current receipts and growth is isomorphic to the joint distribution of current and future receipts: labeling current log receipts r_t, growth $g_t = r_{t+1} - r_t$, and the notch in future income *notch* implies that conditional on current receipts there is a unique level of growth ($g_t = notch - \log r_t$) that puts receipts at the notch in the next year. Bunching will manifest in the distribution of growth from current income as an excess share of organizations growing at rates just below that which brings them to the notch and a reduced share growing to just above it. Future bunching can be estimated conditional on any level of current receipts by constructing a counterfactual distribution of conditional growth rates.

Distortions in the distribution of growth rates can be identified if this distribution does not change sharply as current income varies. Figure 7 conveys the idea behind the identification strategy in this form of dynamic bunching estimation. Panel A shows the distribution of income in the next year for three illustrative ranges of current income. Each conditional distribution of future receipts is centered around the level of current receipts. For each group, the distribution of future income is distorted around the notch, with excess mass just to the left and reduced mass just to the right. Panel B shows the distribution of growth rates for each group, a simple translation of the group's future income. Charities with different levels of current receipts have similarly-shaped growth distributions, except that each has a bunching distortion wherever the notch lies in its distribution (in bold). Local responses to the notch will not affect the growth distribution away from the notch. Each group's growth distribution has a similar shape for most levels of growth and distortions at different levels of the growth than other groups. The extent of the distortions can therefore be estimated by comparing the shape of one group's growth distribution around its notch to the corresponding, undistorted section of the growth distribution among charities that are a different distance from the notch.

The dynamic approach estimates the same measures of income distortions as the static estimates but relies on arguably more attractive assumptions of smoothness of conditional rather than unconditional distributions. The static approach assumes that (1) the distribution of current receipts would be continuous and smooth in the absence of the notch and that (2) deviations from this counterfactual distribution all occur within a neighborhood of the notch. What is required for consistent dynamic estimation of bunching using my approach is that ($1'$) in the absence of the notch the distribution of growth conditional on current receipts would vary smoothly with growth and current receipts, so that one can use the distribution of growth at other income levels as a counterfactual, and that ($2'$) the manipulations of conditional growth all occur within a neighborhood of the notch. It is possible for either pair of assumptions to hold while the other fails, though plausible behavioral patterns seem more likely to violate the static assumptions. In Appendix A I discuss examples of behavior that would yield biased static estimates but consistent dynamic estimates.

I implement the dynamic estimation strategy in two ways. First, in Section 5, I construct bins of the joint distribution of current (log) receipts and conditional growth to the next year's (log) receipts and perform OLS and IV regressions. This reduced-form procedure provides a novel but intuitive means of comparing "treated" charities approaching the notch to "control" charities with similar growth rates but different starting points and hence future receipts away from the notch. With the reduced-form regressions I describe distortions close to the notch and test for heterogeneous responses and long-run effects. In Section 6 I use maximum likelihood to obtain precise estimates of the extent of income manipulation and extensive-margin responses.

5 Dynamic Estimation of Bunching Characteristics

Using a dynamic approach to bunching estimation I present an easily-implemented method for characterizing the manipulation of income around a notch and testing the importance of various factors that might relate to this manipulation. I estimate the propensity to bunch within a neighborhood of the notch, the amount by which income is manipulated within this neighborhood, and the traits that predict whether a charity will bunch.

5.1 Methodology for Describing Bunching

Dynamic estimation provides an opportunity to describe the propensity to bunch, the means by which agents bunch, and heterogeneity in each. The idea is to compare those whose growth will bring them near a notch to a counterfactual constructed using those with similar growth rates but different levels of current income. Here I implement this approach by binning the bivariate distribution of current receipts and growth rates, which allows for convenient graphical illustration and analysis of heterogeneity. The binning approach in this section will provide transparent evidence of whether, conditional on current income, those approaching the notch will move to lower income levels than predicted in order to stay below the notch. In this framework it is straightforward to test which traits predict manipulation of future income.

I now argue that if responses to a notch are local then it is possible to construct treatment groups and control groups that are not selected on the basis of whether they bunch. Consider a range of growth rates, say growth of 10 to 20 percent. Charities with current receipts 20 percent below the notch will be especially likely to have growth in this range, which would put them in the bunching range in the next year. Charities 10 percent below the notch will be especially unlikely to have growth in this range, which would put them in the reduced range in the next year. The group of charities growing 10 to 20 percent from either of these starting points will therefore contain a selected group with too many or too few bunchers. However, there is an intermediate range of current receipts for which the growth range of 10 to 20 percent gives a range of future receipts that spans the notch. Call charities in this intermediate range of current receipts the treatment group. If responses to the notch are local then, for charities with current receipts in the treatment group, the growth range of 10 to 20 percent will include both bunchers and nonbunchers. That is, if we choose a wide enough range of growth rates to include both the bunching and reduced regions then the bunching response should not affect the total share of organizations growing at a rate within this range. For different ranges of growth rates we can identify different treatment groups that grow to a point *near* the notch and construct estimate counterfactual growth rates within this range using organizations with growth in the same range but higher or lower levels of current receipts.

Details of my implementation of the dynamic OLS estimates appear in Appendix A. The goal is to include a set of controls in current receipts and growth that will provide an accurate counterfactual for the group approaching the notch. Here I will simply provide visual evidence that my construction of treatment and control groups is reasonable. Figure 8 displays the probability of being in a particular growth range (growth of log receipts by .1 to .2) as a function of current receipts. The filled circle with confidence intervals indicates the probability among the treatment group of charities for whom this growth rate puts them near the notch in the next year. The probability of being *somewhere* in this growth bin is not distorted for these charities because the bins are wide enough to include both the bunching range and the reduced range. As expected, the share in the growth bin *is* distorted for charities just to either side of the "Near Notch" group (represented by light gray markers in Figure 8) because growth in this range puts these observations squarely in one of the distorted regions on either side of the notch. These charities are excluded from estimation of the counterfactual. These results can be shown for any growth bin x. I define $near\ notch_{xit}$ as an indicator for the treatment group of charities in this growth bin that will be near the notch in the next period and $near\ notch_{it} = \sum_x \left(near\ notch_{xit} * 1_{\{x \leq growth_{it} < x+.1\}}\right)$ for treated charities in any growth bin.

With this strategy of identifying treatment and control groups it remains only to define the outcomes of interest that will characterize responses. The primary outcome in which we should see responses is receipts growth, which should be reduced among charities nearing the notch. One can also examine the growth of other financial variables, including total revenues, expenses, and assets, to determine whether these are affected along with income. Finally, it will be useful to construct an indicator $cross_{xit}$ for growth above the observation-specific rate corresponding to crossing the notch. For each observation in the treatment bin, the growth rate that will bring it to the notch is a simple function of location in the bin.[17] This same function can be applied to all observations, regardless of bin, to obtain the growth rate that *would* correspond to the notch if the observation were in the treatment bin. For growth rate range x to $(x + .1)$, I define $cross_{xit}$ as an indicator for whether the charity grows by more than this rate. I will say an observation "crosses" if $cross_{xit} = 1$. Because $cross_{xit}$ only has significance for the treatment bin, the relative probability of $cross_{xit}$ will be reduced in the bin of interest by the share of charities that bunch and not affected for other bins. As with the treatment variable $near\ notch_{it}$ we can stack regressions for all growth rate ranges if we define $cross_{it} = \sum_x \left(cross_{xit} * 1_{\{x \leq growth_{it} < x+.1\}}\right)$. We can also examine long-run effects by defining $cross_{it+s}$ using the growth rate over the next s years and estimate whether the probability of being across the notch s years is reduced in the future. Finally, interactions of $near\ notch_{it}$ with other variables describing a charity offer straightforward tests for heterogeneity in the bunching response.

5.2 Estimation and Results Describing the Bunching of Charities

I now employ the dynamic design to describe the bunching of charities. Measures of income manipulation are highly significant for charities moving to a bin surrounding the notch. Manipulation occurs only among those not already filing Form 990 and is less common among larger charities. Short-run income manipulation by charities with administrative staff provides suggestive evidence of avoidance behavior, but the notch also

[17] For a growth rate range of x to $(x + .1)$ and bunching range of width $Bwidth$, the treatment bin has minimum value $binmin_{it} = notch - Bwidth - x$. Observations with receipts at the minimum of this bin will cross the notch if they grow by $x + Bwidth$. Other charities in the bin will cross the notch if growth is greater than $x + Bwidth - (r_{it} - binmin_{it})$. $cross_{xit} = 1_{\{growth_{it} > x + Bwidth - (r_{it} - binmin_{it})\}}$.

has long-run effects on growth.

Before presenting regression results using multiple growth bins I provide a visual example using charities growing by .1 to .2 log points. Figure 9 plots the constructed variable $cross_{.1it}$ as a function of current income. The filled circle with standard error bands shows the share that cross the notch among those for which $near\ notch_{.1it} = 1$, while empty circles and the quadratic fit display the outcome for charities with higher or lower current receipts. The counterfactual implies that over 40 percent of the charities nearing the notch should have crossed it, but instead less than 35 percent do so. These findings extend to charities in the other growth rate ranges that I now combine in estimation.

Regression analysis shows highly significant manipulation of receipts. In column (1) of Table 5 we see that receipt growth of charities that near the notch is lowered by .0017 log points. The average reduction is therefore about 0.17% of \$100,000, or \$170. The average is taken over all charities nearing the notch, whether they bunch or not. In column (2) we see that the probability of achieving growth that would imply crossing the notch is reduced by 4 percentage points. This regression of $cross_{it}$ on $near\ notch_{it}$ and controls is also the first stage of an instrumental variables estimate of receipt manipulation by bunchers themselves, the second stage of which is presented in column (3). The identifying assumption of the IV specification is that receipt growth of charities in the group approaching the notch only deviates from the counterfactual due to their responses to keep receipts below the notch. The IV results show that the average buncher reduces reported receipts by .0423 log points, or about \$4500. Unfortunately, the expense and asset growth outcomes that might signal the extent of avoidance are not precisely estimated; standard errors are larger than the direct effect of the notch on receipts, and underlying growth rates for these variables are similar to that of receipts. Such regressions may prove more informative in settings where more data are available or growth rates are less variable.

Next I estimate the effect of the notch on long-run growth. Table 6 displays the results of 12 regressions for the probability of crossing the notch in t years. The results show that the notch reduces crossing by about 1.5 percentage points for over a decade. The reduction in crossing is relative to the counterfactual share that should cross (not shown), which grows from 40 percent in year one to a bit over 75 percent in year ten.

In addition to these average responses, the dynamic estimation strategy provides illuminating tests for heterogeneity in responsiveness. Table 7 shows that smaller organizations are more likely to reduce income to stay below the notch. The outcome for each regression is the indicator variable $cross_{it}$. Interactions of total revenue, expenses, and assets (all in logs) with $near\ notch_{it}$ reveal that larger charities are more likely to cross the notch when approaching it, i.e. less likely to reduce income to avoid crossing. The magnitude of the coefficients implies that a one percent increase in a charity's expenses or assets is associated with about a 1.5 percentage point (2.5 percent) reduction in the probability of manipulating receipts when approaching the notch in the next year. Including all of these variables and their interactions eliminates the predictive power of total revenue but leaves expenses and assets as highly significant determinants of bunching. The fact that large organizations are less likely to bunch supports the idea that the long form imposes administrative expenses, some of which are likely related to transitioning to an accounting infrastructure that facilitates detailed financial reporting.

The first four columns of Table 7 show that size is predictive of income manipulation but may have more than one interpretation. Since most years of the NCCS data do not include a variable indicating which form was filed, it could be that larger organizations respond less because they are already filing Form 990 or because they adopt the form more quickly when reaching the notch, regardless of which form they filed before. To address this questions I incorporate data from the IRS Statistics of Income files for a random sample of 990-EZ filers. Columns (5) and (6) report results of regressions that only include observations moving to the notch if they appear in the IRS Statistics of Income 990-EZ sample. Column (5) of Table 7 shows that 990-EZ filers are less likely to cross the notch, consistent with an adjustment cost. The interaction terms in column (6) are no longer significant due to the reduced sample of organizations nearing the notch, but the point estimates are quite similar to those in other columns. It appears, therefore, that large charities are not just more likely to file Form 990 before required but are less likely to manipulate income to stay below the notch even if they previously filed Form 990-EZ.

The final dimension of heterogeneity for which I present results is staffing. Charities with paid staff may be less willing to file Form 990 and more able to manipulate income to avoid filing the longer form. Unfortunately, the data do not include the staffing line item for charities filing Form 990-EZ. To examine heterogeneity by future staffing I restrict attention to charities that have receipts above the notch at some point in the sample. The data include Form 990 staffing variables "Compensation" (for officers and directors), "Other Salary" (for others), and "Payroll Taxes."[18] The regression results in Table 8 reveal how staffing variables and their interactions with $near\ notch_{it}$ predict manipulation according to the outcome $cross_{it}$. Charities with paid administrative staff, whether measured by "Other Salary" or "Payroll Taxes," are less likely to cross the notch when they first approach it. This result provides suggestive evidence that while the notch was found to have permanent effects on some charities' growth it also leads to some temporary avoidance. I deem these results "suggestive" because the notch was shown to have permanent effects on the share crossing, which implies that the sample of charities that eventually cross may be selected based on characteristics related to the staffing variables.

[18]Just under half of the estimation sample has "Other Salary" when above the notch, and median Other Salary is between $30,000 and $35,000.

A few other covariates suggest variation in the incentives or ability to bunch. These results are available by request. First, if assets are above $250,000 then the organization must file Form 990 regardless of receipts level. Only charities below the asset notch would be expected to bunch, and this is confirmed in regression analysis (though this does not eliminate the size effects presented in Table 7). Other financial variables that appear with the same wording on both forms include fundraising event income and inventory sales, neither of which predicts bunching. It is also possible to test for disclosure costs, albeit imperfectly, by examining whether some information that appears on Form 990 and not Form 990-EZ predicts bunching. As with the staffing variables, these measures must be defined in years that charities file the long form. Using the value of each variable in the first year after a charity crosses the notch, I find no evidence that charities avoid filing in order to conceal fundraising expenses or sources of business income unrelated to the charitable purpose. The cost of disclosing other variables appearing on the long form, including personal benefit contracts and controlled entities, could not be tested because these variables are not captured in the data.

In summary of the OLS and IV results, I find significant manipulation of income when nearing the notch. Consistent with adjustment costs, large charities and those that filed Form 990 previously are less likely to avoid being above the notch. Short-term manipulation by charities with administrative staff suggests avoidance, but the notch also has significant effects on growth in the long run. I do not find evidence that charities reduce income to avoid disclosing other information but do not have sufficient data to completely rule out this possibility.

6 Dynamic Estimation of the Quantity of Bunching

Having described characteristics of bunching, I now turn to estimation of the extent of bunching. As before, the goal is to quantify distortions in the joint distribution of current and future receipts. I construct a maximum likelihood estimator that identifies the entire counterfactual distribution of growth conditional on current receipts by comparing observations that differ in current receipts and hence in the level of growth that would bring them to the notch. Bunching is estimated as the difference between the observed share moving to a region just below the notch and the share predicted by the counterfactual distribution of conditional growth. The approach can account for notch-related attrition due to extensive margin responses or other features specific to the empirical setting.

6.1 Methodology of Dynamic Estimation of the Quantity of Bunching

Distortions in the joint distribution of an agent's income in different years identify bunching in dynamic settings. While the tools presented in the previous section can provide estimates of the quantity of bunching, a maximum likelihood approach offers advantages in terms of efficiency and clarity of assumptions. For example, it is straightforward to define parts of the growth distribution from which agents may not be observed in the next year's data and test for these systematic deviations from a random sample. The counterfactual conditional growth distribution is estimated as a smooth but flexible function with form that is identified by observations throughout the distribution of current receipts.

Maximum likelihood estimation provides an efficient way to perform joint estimation of the prevalence of attrition and income responses. Dynamic OLS bunching estimates offer potential advantages over static estimates but retain some drawbacks. Without further adjustments, OLS may not provide a consistent estimate of the bunching propensity if organizations are more likely to go missing from the data when their receipts exceed the notch. Even if they are consistent, the OLS estimates may not be efficient because binning the data by current receipts and growth rate treats observations within a bin as equivalent, and the choice of bin widths and locations is necessarily ad-hoc when the data are continuous. These issues can be addressed with a maximum likelihood estimator.

The reasoning behind the MLE approach is the same as that for the OLS estimator: the level of growth that will take an organization to the Form 990 notch depends on current receipts. It would be possible to estimate the entire joint distribution of current income and future income, but focusing on the distribution of growth conditional on current receipts provides computational benefits and isolates the desired variation in the growth distribution rather than trying to simultaneously recover the static income distribution. I therefore estimate the conditional cdf of counterfactual growth, $F(g|r)$.[19]

The maximum likelihood estimate can be implemented by defining and estimating the parameters of a flexible function for the conditional distribution of growth. Growth g is defined as the change in an organization's log receipts from the current value of r to its value one year later. I first define the latent cdf $F(g|r)$ that would be observed if no observations were bunching or going missing, then incorporate these responses into the distribution $F^*(g|r)$ that is fit to the data. To parametrize the counterfactual growth distribution I assume it falls within a flexible class of widely-used functions. Because the data have fat tails and a kink at zero growth, the Laplace distribution provides a natural choice. Laplace (or "double exponential") distributions have been used extensively to model financial data and "are rapidly becoming distributions of first choice whenever 'something' with heavier than Gaussian tails is observed in the data" (Kotz et al., 2001).[20] The Laplace distribution describes the difference between two independent

[19] The joint density of growth and receipts can be estimated by maximizing $\sum_i \log(f(g_i, r_i)) = \sum_i \log(f_g(g_i|r_i) \cdot f_r(r_i)) = \sum_i [\log(f_g(g_i|r_i)) + \log(f_r(r_i))] = \sum_i \log(f_r(r_i)) + \sum_i \log(f_g(g_i|r_i))$. Maximizing only the second term will provide consistent estimates of the parameters of the conditional growth density but may be less efficient than maximizing the joint density if all years are included such that current receipts in one year also enter the growth rate for the previous year.

[20] See Kozubowski and Nadarajah (2010) for other recent applications.

exponentially-distributed random variables. While the distribution of charities' reported receipts appears to be approximately exponential, an organization's future receipts are certainly not independent of current receipts, providing one important reason to allow for flexibility in the conditional growth distribution. I estimate a modification of the Laplace cdf by allowing for flexible functions $P_l(g, r, \theta)$ and $P_u(g, r, \theta)$ to enter the lower and upper pieces of the distribution:[21]

$$F(g|r) = \begin{cases} \exp\left(P_l(g, r, \theta)\right) & g < \theta(r) \\ 1 - \exp\left(P_u(g, r, \theta)\right) & g \geq \theta(r) \end{cases}$$

I describe the main points of the maximum likelihood estimation here and provide details in Appendix C. Using the latent density of log receipt growth I then specify the form of the observed density to account for bunching and for missing data. To adjust for bunching I define a conditional omitted region of growth rates that take each current level of receipts to a region around the notch. I count observations that grow to the omitted region so that they are not treated as missing, but I exclude them from estimation of the shape of the latent distribution. Next I account for attrition, which could be due to late filing, earning receipts below the level at which filing is required, shutting down, merging, or simply non-compliance. I estimate three types of attrition. First, I include terms that are constant or linear in current receipts to capture basic, random attrition. Second, I adjust the observed conditional growth densities to account for truncation of the sample due to the fact that organizations with receipts below \$25,000 do not have to report. Third, I use the latent cdf to determine the share of observations that should cross the notch from each level of current receipts and allow a heightened probability that charities that should cross the notch instead go missing. This last parameter estimates the extra share of organizations that go missing because the requirement to file the longer form for the first time induces late filing or nonfiling. This last parameter is identified by the way in which attrition varies with current receipts, because the share that should cross the notch increases as current receipts approach the notch.

I will report estimates describing the bunching response and missing observations. The first parameter of interest is the bunching propensity or bunching share (b in the discussion of theory Section 2), which reveals the share of bunchers among the total number that should move to the reduced region. is identified by comparing the observed distribution of growth rates to the counterfactual distribution. I allow this bunching parameter to take a different value for charities coming from below the notch than for charities already above the notch. I use the estimated bunching shares and the counterfactual distribution of growth to calculate the excess mass that is observed in the bunching range in the next year and the reduction in the mass above

[21] The symmetric Laplace distribution with location parameter θ and scale parameter σ has this form with $P_l(g, r, \theta) = P_u(g, r, \theta) = \left|\frac{g-\theta}{\sigma}\right| - \log(2)$.

the notch. I also report estimates of excess attrition among those below the notch and those crossing it, where the latter is identified by variation across current receipts in the counterfactual share with growth that would put them above the notch.

6.2 Results of Dynamic Estimation of the Quantity of Bunching

Now I turn to the maximum likelihood estimates of the amount of bunching. Results are broadly consistent with the bunching measures obtained from static estimation and the ordinary least squares dynamic estimation. Roughly ten percent of the charities that should cross the notch by a small amount in each year will instead remain below it. Extensive-margin responses are significant and explain the difference between the static estimates of the excess mass below the notch and the reduced mass above the notch.

Table 9 displays the results of MLE estimation of the extent of bunching and systematic attrition. The first parameter estimate in each column gives the bunching propensity among charities that have current receipts below the notch. In the basic specification, 9.3 percent of such charities that should have future receipts just above the notch will instead reduce reported receipts to stay below the notch. The second row shows the bunching propensity for those with current receipts above the notch, which is always estimated to be less than 0.14 percent and never significantly different from zero. Charities coming from above have already filed Form 990 and have less incentive to bunch if the marginal cost of filing is largely a one-time adjustment cost. The lack of bunching by charities coming from above the notch is consistent with (unreported) results from the reduced-form estimation strategy of Section 4.

Attrition is significantly related to current receipts. Columns (1) through (3) of Table 10 display results for different specifications of attrition as a function of current and future receipts. Adding a simple constant term for charities currently below the notch as in column (2) reveals that these charities are about 1.3 percentage points more likely to go missing in the next year but has little effect on the bunching propensity estimates. The estimated bunching propensity decreases, however, when allowing for greater attrition among charities that would have crossed the notch, as in column 3. Failure to account for these extensive margin responses leaves only the bunching estimate to account for the full reduction of mass above the notch, whereas the flexible approach distinguishes the response of manipulating income from the response of leaving the sample. The flexible model minimizes the Akaike Information Criterion, reinforcing the importance of systematic attrition.

The final rows of Table 9 reveal the estimated excess share of charities below the notch and reduction in thee share above it. The excess and reduction are found by aggregating the bunching and attrition propensities across all observations according to their counterfactual probability of moving to the reduced region. In the baseline specification, .183 percent of all organizations will manipulate income to bunch in the next year. By construction, the excess and reduced mass are equal. The estimate of .183 lies in between the static estimates of the excess and the reduction, as reported in the last column. Allowing for excess attrition below the notch has a minor impact on the excess estimate, but allowing for extensive-margin responses gives estimates quite similar to those of the static approach. The dynamic approach therefore confirms the static estimate that about .15 percent of charities manipulate income and provides evidence that the additional .1 percent reduction in the mass of charities above the notch is due to extensive-margin responses.

Table 10 shows robustness of the baseline dynamic estimates to the choice of width for the bunching and reduced ranges. The baseline dynamic regression, with an omitted region of $80-130,000, corresponds to column (3) in Table 9. The dynamic estimates vary with the omitted range as one would expect. If the researcher overly restricts the omitted range so that it does not cover the full range over which income is manipulated then bunching will be underestimated. Accordingly, the dynamic estimates capture more bunching as the reduced range used in the estimation is expanded from $10,000 in width to about $30,000. Further widening does not affect the estimate much because all bunching has been captured. The static estimates are also fairly robust for sufficiently wide omitted ranges. The main difference between the patterns of dynamic and static estimates is that the static estimate of the excess grows rather than shrinks when the reduced range is overly restricted because observations in the true reduced range are used to estimate the counterfactual, causing it to be biased downwards. In practice there should be no problem so long as the reduced range is made sufficiently wide.

7 Conclusion

This study provides new evidence on the behavior of charities and responses to threshold policies. The IRS income threshold for filing simplified returns produces a clear distortion in the distribution of reported income. The average charity is willing to reduce income by $600 to $1000 to avoid reporting more information, but this average masks considerable heterogeneity, and the fact that small charities who had previously filed the simplified form were most likely to manipulate income provides evidence that much of the compliance cost is a one-time adjustment. Responses appear to consist of both short-run manipulation of income and permanent distortions of income growth.

The results of this study highlight several benefits of incorporating dynamics into bunching estimation. Conditioning on past income provides a different identification strategy and provides new opportunities to describe behavior by estimating extensive-margin responses, preference heterogeneity, long-run effects, and the extent of avoidance. Dynamic bunching techniques could be used to analyze responses to thresholds in many other settings, including social welfare programs with income-eligibility limits, tolls and security checkpoints, product pricing with quantity discounts, or rewards programs for charitable giving.

References

Acemoglu, Daron and Joshua D. Angrist, "Consequences of Employment Protection? The Case of the Americans With Disabilities Act," *Journal of Political Economy*, 2001, *109(5)*, 915–957.

Andreoni, James, "Philanthropy," *Handbook on the Economics of Giving, Reciprocity and Altruism*, 2006, *1*, 1201–1269.

Angelini, Paolo and Andrea Generale, "On the Evolution of Firm Size Distributions," *American Economic Review*, 2008, *98(1)*, 426–438.

Bastani, Spencer and Håkan Selin, "Bunching and Non-Bunching at Kink Points of the Swedish Tax Schedule," *CESifo Working Paper Series No. 3865*, 2012.

Blumenthal, Marsha and Laura Kalambokidis, "The Compliance Costs of Maintaining Tax Exempt Status," *National Tax Journal*, 2006, *56*, 235–52.

Cabral, Luis M. B. and José Mata, "On the Evolution of the Firm Size Distribution: Facts and Theory," *American Economic Review*, 2003, *93(4)*, 1075–90.

Carillo, Paul E., M. Shahe Emran, and Anita Rivadeneira, "Do Cheaters Bunch Together? Profit Taxes, Withholding Rates and Tax Evasion," April 2012.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *The Quarterly Journal of Economics*, 2011, *126*, 749–804.

DellaVigna, Stefano, James A. List, and Ulrike Malmendier, "Testing for Altruism and Social Pressure in Charitable Giving," *The Quarterly Journal of Economics*, 2012, *127(1)*, 1–56.

Dharmapala, Dhammika, Joel Slemrod, and John Douglas Wilson, "Tax policy and the missing middle: Optimal tax remittance with firm-level administrative costs," *Journal of Public Economics*, 2011, *95*, 1036–1047.

Diamond, Peter, "Optimal tax treatment of private contributions for public goods with and without warm glow preferences," *Journal of Public Economics*, 2006, *90* (4-5), 897–919.

Dixon, Peter B., Mark R. Picton, and Maureen T. Rimmer, "Payroll Taxes: Thresholds, Firm Sizes, Dead-Weight Losses and Commonwealth Grants Commission Funding," *The Economic Record*, 2004, *80(250)*, 289–301.

Froelich, Karen A. and Terry W. Knoepfle, "Internal Revenue Service 990 Data: Fact or Fiction?," *Nonprofit and Voluntary Sector Quarterly*, 1996, *25(1)*, 40–52.

Gelber, Alexander, Damon Jones, and Dan Sacks, "Individual Earnings Adjustment to Policy: Evidence from the Social Security Earnings Test," 2012. Wharton School of the University of Pennsylvania, mimeo.

Greenstone, Michael, Paul Oyer, and Annette Vissing-Jorgensen, "Mandated Disclosure, Stock Returns, and the 1964 Securities Acts Amendments," *The Quarterly Journal of Economics*, 2006, *121* (2), pp. 399–460.

Iliev, Peter, "The Effect of SOX Section 404: Costs, Earnings Quality, and Stock Price," *The Journal of Finance*, 2010, *65*, 1163–1196.

Internal Revenue Service, "IRS Data Book, Publication 55b," Technical Report 1998-2005.

_ , "2007 Instructions for Form 990 and Form 990-EZ," 2007.

Keen, Michael and Jack Mintz, "The Optimal Threshold for a Value-Added Tax," *Journal of Pub*, 2004, *88*, 559–576.

Kleven, Henrik J. and Mazhar Waseem, "Behavioral Responses to Notches: Evidence from Pakistani Tax Records," September 2012.

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

Kopczuk, Wojciech and David Munroe, "Mansion Tax: The Effect of Transfer Taxes On Residential Real Estate Market," 2012. Columbia University, mimeo.

Kotz, Samuel, Tomasz J. Kozubowski, and Krzysztof Podgórski, *The Laplace Distribution and Generalizations: A Revisit With Applications To Communications, Economics, Engineering, and Finance*, Birkhäuser, 2001.

Kozubowski, Tomasz J. and Saralees Nadarajah, "Multitude of Laplace Distributions," *Statistical Papers*, 2010, *51(1)*, 127–48.

Manoli, Dayanand and Andrea Weber, "Nonparametric Evidence of the Effects of Financial Incentives on Retirement Decisions," *NBER Working Paper No. 17320*, 2011.

Marx, Benjamin M., "Effects of Regulation on Donations to Charitable Foundations," 2012. Columbia University, mimeo.

National Center for Charitable Statistics, "Guide to Using NCCS Data," August 2006.

Onji, Kazuki, "The Response of Firms to Eligibility Thresholds: Evidence from the Japanese Value-Added Tax," *Journal of Public Economics*, 2009, *93*, 766–775.

Ramnath, Shanthi, "Taxpayers' Response to Notches: Evidence from the Saver's Credit," 2012. University of Michigan, mimeo.

Roeger, Katie L., Amy Blackwood, and Sarah L. Pettijohn, "The Nonprofit Almanac 2012," Technical Report, Urban Institute Press 2012.

Saez, Emmanuel, "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *The Quarterly Journal of Economics*, 2002, *117* (3), pp. 1039–1073.

_ , "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2010, *2(3)*, 180–212.

Sallee, James M. and Joel Slemrod, "Car Notches: Strategic Automaker Responses to Fuel Economy Policy," *NBER working paper 16604*, 2010.

Schivardi, Fabiano and Roberto Torrini, "Identifying the Effects of Firing Restrictions Through Size-Contingent Differences in Regulation," *Labour Economics*, 2007, *15*, 482–511.

Slemrod, Joel, "A General Model of the Behavioral Response to Taxation," *International Tax and Public Finance*, 2001, *8(2)*, 119–128.

_ , "Buenas Notches: Lines and Notches in Tax System Design," 2010. University of Michigan, mimeo.

_ **and Marsha Blumenthal**, "Measuring Taxpayer Burden and Attitudes for Large Corporations," Technical Report, Report to the Coordinated Examination Program of the Internal Revenue Service 1993.

_ **and Varsha Venkatesh**, "The Income Tax Compliance Costs of Large and Mid-Sized Businesses," Technical Report, Report to the Internal Revenue SErvice Large and Mid-Size Business Division 2002.

Weber, Caroline E., "Does the Earned Income Tax Credit Reduce Saving by Low-Income Households?," April 2012.

Yelowitz, Aaron S., "The Medicaid Notch, Labor Supply, and Welfare Participation: Evidence from Eligibility Expansions," *The Quarterly Journal of Economics*, 1995, *110* (4), pp. 909–939.

Appendix A - Discussion of the Relationship Between Dynamic and Static Bunching Estimates

Dynamic bunching estimation offers several benefits as a complement to static estimation in settings for which panel data is available. The dynamic estimates can illuminate the nature of heterogeneity in responses, as seen in Section 4, and can test for attrition that is endogenous to a notch, as in Section 5. Here I argue that, in addition, the identifying assumptions of dynamic bunching estimation are arguably more plausible than those used in static estimation, particularly for settings where agents face a notch repeatedly.

Repeated bunching may lead to a violation of the identifying assumptions of the static approach. Consider, for example, agents that grow income at a constant rate every year unless they approach a notch, in which case they never cross. In the current setting, this might describe a charity with a low growth rate and a high discount factor. Even if the charity is fully forward-looking, it may not deem future growth to be of sufficient value to compensate for the cost of crossing the notch. Conditional on current income, then, the charity makes a rational decision of the kind motivating bunching analysis, in which bunching in the next period is preferred to income in a range just above the notch. In the long-run, however, the constant growth this charity would have achieved in the absence of the notch would raise its income far above the notch. In this case, the static counterfactual will underestimate the number of agents that should be above the notch no matter how wide the researcher allows the omitted region to be. If the distribution is only affected above the notch then it might be possible to estimate the counterfactual using only observations below the notch and projecting the results to higher incomes. However, projection may be unreliable in practice, as was shown in Table 3, and the distribution may also be affected below the notch if agents that get "stuck" at the notch then experience negative income shocks.

Income that exhibits a high degree of serial correlation could also pose a concern for static bunching estimates. To see that serial correlation may violate the smoothness assumption, consider the extreme case in which the conditional distribution of income in the next period is discrete-continuous with strictly positive mass at today's income level. Say that income has observed distribution $f_t(y_t)$ in the current year, and the pre-bunching (counterfactual) cumulative distribution function (cdf) for the following year is given by $F_{t+1}(y_{t+1}) = \int G_{t+1}(y_{t+1}|y_t) f_t(y_t)\, dy$, with $G_t(y_{t+1}|y_t) = \alpha \cdot 1_{\{y_{t+1} \geq y_t\}} + (1-\alpha) H_t(y_{t+1}|y_t)$ for some constant $\alpha \in (0,1)$ and continuous cdf $H_t(y_{t+1}|y_t)$. Say there is a notch at $y_t = n$ and bunching at the notch in current year. Then

$$\lim_{y_{t+1} \to n^+} F_{t+1}(y_{t+1}) - \lim_{y_{t+1} \to n^-} F_{t+1}(y_{t+1})$$

$$= \lim_{y_{t+1} \to n^+} \int \left[\alpha \cdot 1_{\{y_{t+1} \geq y_t\}} + (1-\alpha) H_t (y_{t+1}|y_t) \right] f_t (y_t) \, dy$$

$$- \lim_{y_{t+1} \to n^-} \int \left[\alpha \cdot 1_{\{y_{t+1} \geq y_t\}} + (1-\alpha) H_t (y_{t+1}|y_t) \right] f_t (y_t) \, dy$$

$$= \alpha \left[\lim_{y_{t+1} \to n^+} F_t (y_{t+1}) - \lim_{y_{t+1} \to n^-} F_t (y_{t+1}) \right]$$

$$\neq 0$$

The difference between these limits is not zero because current bunching implies the current income distribution is discontinuous at the notch. Because income is highly persistent, the discontinuity at the notch will remain in the future even without further bunching. To create problems in practice, the distribution of growth need not truly be discrete-continuous, but simply concentrated around a particular growth rate (such as zero).

Figure A.1 provides an illustration of how serially correlated income could affect estimation. The figure shows projected counterfactuals using the static methodology and the methodology presented in Section 6. The dashed line depicting the dynamic estimate of the counterfactual distribution coincides nearly exactly with the dotted line depicting the static estimate. The similarity of the dynamic and static counterfactuals highlights the equivalence of the two approaches in this setting. The solid line, however, describes the counterfactual if 10 percent have zero growth and 90 percent of charities follow the dynamic estimate of the growth distribution conditional on their receipts in the prior year. In this case, even the counterfactual is discontinuous simply because there are more charities below the notch in the prior year. If many agents have income growth close to zero, the income distribution would exhibit bunching even if the notch was removed and agents had no further propensity to bunch. While the notch is in place, mass may accumulate over time in the bunching range, leading to biased estimates of the propensity to bunch in any particular year. These issues may arise, for example, in the context of individual incomes at an inflation-indexed notch if cost-of-living adjustments to wages are also indexed to inflation.

The growth distribution of charitable organizations is sufficiently disperse that the dynamic estimates of the bunching quantity closely follow the static estimates, but there is some evidence the growth distribution has become increasingly distorted around the notch. Annual estimates show that the discontinuity in the density of receipts at the notch has grown steadily over most of the sample period. Among charities in the NCCS data, the overall share with receipts above \$100,000 has steadily declined from over 66% in

1989 to less than 62% in 2002. Using entry rates and a simple binary transition matrix estimated over the full sample, I calculate a stationary distribution in which fewer than 42% of charities have receipts above $100,000. A continuation of this trend might have lead to sufficient accumulation that significant excess mass would remain for some amount of time after the removal of the notch, but there had not been sufficient accumulation by the time the Form 990 notch was moved. Future research can apply the methodology developed in this paper to assess the importance of dynamics in other settings.

Appendix B - Details of Dynamic Ordinary Least Squares Estimation

Section 5 introduced a reduced-form approach to dynamic bunching estimation to characterize heterogeneity and long-run effects. This appendix details the implementation in this paper, including the estimating equation, sample selection and bin construction, choice and test of omitted range, an instrumental variables specification, and a test for long-run effects.

Estimating outcome Y_{it+1} in the dynamic ordinary least squares approach involves stacking multiple growth rate ranges and estimating equations of the form

$$(2) \quad Y_{it+1} = \beta \cdot near\ notch_{it} + \sum_{j=1}^{J} \alpha_j r_{it}^{j} + \sum_{k=0}^{K} \sum_{a=1}^{A} \gamma_{ka} r_{it}^{k} D\left(f\left(a\right), f\left(a+1\right)\right)_{it+1}$$

where r_{it} is current recentered log gross receipts and $D\left(f\left(a\right), f\left(a+1\right)\right)_{it+1} = 1\left\{r_{it+1} - r_{it} \in \left[f\left(a\right), f\left(a+1\right)\right)\right\}$ is an indicator for growth falling within a particular range. The estimating equation allows for specifying greater or lesser flexibility in the controls, as desired. The double sum contains an interaction term that allows for a separate pattern of variation across receipts within each growth rate range. The expression encompasses the simple case of growth rate dummies not interacted with current receipts (K=0). Results are highly robust to different specifications. I present specifications with a full set of growth range dummies and the interactions of each growth rate bin with a quadratic function of receipts (so that $J = 2$ and $K = 2$ in the estimating equation). Potential outcomes of interest include growth of receipts, expenses, and assets. Bunching would imply that receipts would grow by less among the $near\ notch_{it}$ treatment group. If this reduction represents real income losses (rather than avoidance) then this group should exhibit concurrent reductions in expenses or assets.

Throughout I use the sample of observations with $r_{it} \in [-1, 2)$ and consider growth rate ranges of the form $[x, x + .1)$ with $x \in [0, .9)$. I show results only for positive growth rates because, as shown in Section 6, essentially all responses are due to charities with current incomes below the notch. Similar results obtain when excluding the bin of lowest growth rates ($x = 0$), for which $near notch_{xit}$ indicates charities with current receipts already in the neighborhood of the omitted region. Statistical testing confirms the visual evidence that observations are binned in such a way that that there is no net distortion to the share of charities in the treatment bins moving to a neighborhood of the notch. Running a regression for each growth rate range, I perform a Wald test of the hypothesis that $\forall x, near\ notch_{xit}$ has no effect on the probability of growth in range x. The test fails to reject, with p value .1361. That the probability of growing to the omitted region is not significantly different from the counterfactual provides evidence that the specified omitted region includes a sufficient range to include organizations whether or not they bunch.

As described in the text, $near notch_{it}$ is a dummy for charities moving to an omitted region that straddles the notch. I present results for current log receipt bins of width .05, log growth bins of width .1, and an omitted region of $r_{it+1} \in [-.08, .07)$. Estimates are qualitatively similar when using receipt bins of width .03 or .1 and growth rate bins of width .05 or .15. Widening the omitted region increases the number of bunchers in the treatment group, which should increase the precision of estimated responses. Because responses are local, widening the omitted region increases the total number in the group by even more than it increases the number of included bunchers, so that for a given level of precision it becomes necessary for bunchers' responses to be larger to distinguish the average response from zero. Such tradeoffs suggest an opportunity to develop an econometric procedure for optimally constructing the bins, but I leave this for future research.

Because the regressions estimate an average response among those that respond and those that do not, it is also useful to estimate the amount by which the bunching charities manipulate their income. To calculate the receipt reductions of bunchers one can relate the reductions of receipts that is estimated with equation (2) and relate this reduction to the share of charities that bunch, which is estimated by equation (2) when the outcome is $cross_{it}$. It is natural, then, to perform Two Stage Least Squares estimation with receipt growth as the outcome and $near\ notch_{it}$ as an instrument for $cross_{it}$. The exclusion restriction would require that charities nearing the notch only reduce their receipts in order to stay below the notch, a reasonable assumption given that these charities are spread through the distribution of current receipts and are compared to other charities with growth rates in the same range as theirs. The coefficient on $cross_{it}$ in the second stage provides a measure of bunchers' average reduction of reported income to avoid filing Form 990.

Lastly, tests for long-run effects merit a brief note on sample selection. I examine long-run effects using the outcome $cross_{xit+s}$ for s ranging from 1 to 12. The specification requires that year zero falls in 1997 or earlier so that each organization can be observed for all twelve years. The sample size generally decreases with the horizon as organizations go missing from the data. Restriction of the sample to charities that appear in all twelve subsequent years would reduce the sample by a prohibitive 90 percent.

Appendix C - Details of Dynamic Maximum Likelihood Estimation

Section 6 introduced the dynamic estimation of bunching by maximum likelihood. The details of this approach follow. I describe the observed distribution as a function of the latent distribution and of parameters governing bunching and attrition.

It is possible to perform maximum likelihood estimation by estimating a flexible function for the pdf and constraining it to integrate to unity, but starting from the cdf offers several advantages. First, it is desirable to estimate excess attrition among those who cross above the notch or below the point of sample truncation, and the cdf gives the probabilities of these occurrences. Second, the cdf makes it straightforward to constrain the reduced mass to equal the bunching mass (except for differences due to systematic attrition). Third, truncation requires integration of the likelihood between limits that vary with the level of current receipts, a practical issue for multidimensional integration programs. A disadvantage of specifying the cdf is the need for functions that appear more arbitrary than their derivatives. For example, I include inverse tangents to allow for curvature at growth rates close to zero because the derivative of $\arctan(x)$ is $\frac{1}{1+x^2}$.

The latent cdf of conditional growth is given by

$$F(g|r) = \begin{cases} \exp\left(P_l(g,r,\theta)\right) & g < \theta(r) \\ 1 - \exp\left(P_u(g,r,\theta)\right) & g \geq \theta(r) \end{cases}$$

$$P_l(g,r,\theta) = \pi_0^l + \tau_0^l r + \left(\pi_1^l + \tau_1^l r\right)(g-\theta) + \left(\pi_2^l + \tau_2^l r\right)\left[\exp(g-\theta) - 1\right]$$

$$P_u(g,r,\theta) = \pi_0^u + \tau_0^u r + \left(\pi_1^u + \tau_1^u r\right)(g-\theta) + \left(\pi_2^u + \tau_2^u r\right)\left[\exp\left(-(g-\theta)\right) - 1\right]$$

$$+ \left(\pi_3^u + \tau_3^u r\right)\left[\exp\left(-(g-\theta)^2\right) - 1\right] + \left(\pi_4^u + \tau_4^u r\right)\arctan\left(\left(\pi_4^u + \tau_4^u r\right)(g-\theta)\right)$$

I now list and impose as needed the conditions that ensure $F(g|r)$ is a cdf. First, the function must have infimum 0 and supremum 1. The appropriate limits can be achieved by two restrictions on the parameters:

1. $\left(\pi_1^l + \tau_1^l r\right) < 0 \Rightarrow \lim_{g\to-\infty} P_l(g,r,\theta) = -\infty \Leftrightarrow \lim_{g\to-\infty} F(g|r) = 0$
2. $\left(\pi_1^u + \tau_1^u r\right) < 0 \Rightarrow \lim_{g\to\infty} P_u(g,r,\theta) = -\infty \Leftrightarrow \lim_{g\to\infty} F(g|r) = 1$

Both constraints are easily implemented by using exponentiated coefficients in the numerical maximization.

Second, $F(g|r)$ must be nondecreasing. Because the posited functional form has one point of nondifferentiability at $g = \theta$, the nondecreasing property requires $\lim_{g\to\theta^-} F(g|r) \leq \lim_{g\to\theta^+} F(g|r)$. I require this relation to hold with equality, giving continuity of the cdf and ruling out point mass at zero growth. This gives

$$\exp\left(P_l\left(\theta,r,\theta\right)\right) = 1-\exp\left(P_u\left(\theta,r,\theta\right)\right)$$

$$\exp\left(\pi_0^l+\tau_0^l r\right) = 1-\exp\left(\pi_0^u+\tau_0^u r\right)$$

$$3.\ \pi_0^l+\tau_0^l r = \log\left(1-\exp\left(\pi_0^u+\tau_0^u r\right)\right)$$

The implied latent density is

$$f\left(g|r\right) = \begin{cases} P_l'\left(g,r,\theta\right)\exp\left(P_l\left(g,r,\theta\right)\right) & g<\theta\left(r\right) \\ -P_u'\left(g,r,\theta\right)\exp\left(P_u\left(g,r,\theta\right)\right) & g\geq\theta\left(r\right) \end{cases}$$

where $P_l'\left(g,r,\theta\right) =$ and $P_u'\left(g,r,\theta\right)$ are derivatives with respect to g. These derivatives can be assured of the correct sign by exponentiating each of the relevant coefficients, but this would impose more than is required because nonnegativity of the density does not necessitate that *all* the coefficients have the same sign. Instead I simply impose a prohibitive penalty on the value of the likelihood function if the pdf is negative for any observations. Similarly, I do not impose conditions 1 and 2, which arise naturally during the optimization, but I do impose condition 3, which has the added benefit of reducing the number of parameters to be estimated.

To measure bunching I estimate b, the share of mass from the reduced region that instead appears in the bunching region. I specify a vector for b, allowing the bunching propensity to depend on whether current receipts are above the notch, but in either case require the bunching mass to equal the reduced mass. I define $notch := \log\left(100,000\right)$ as the Form 990 receipts notch and allow organizations to shift receipts from a region of width $Rwidth$ to a region of width $Bwidth$. Thus, there is excess mass B in the bunching region $g+r\in[notch-Bwidth, notch)$ that would otherwise lie in the reduced region $g+r\in[notch, notch+Rwidth)$. Combining these ranges gives an omitted region of $g+r\in[notch-Bwidth, notch+Rwidth)$. I do not use charities moving to the omitted region to identify the shape of the latent distribution. However, I incorporate these observations to estimate bunching and attrition parameters. To do this I generate a variable $g*$ equal to $(notch+Rwidth-r)$ for charities moving to the reduced range, $(notch-r)$ for charities moving to the bunching range, and g for other charities. The fact that $g*$ is assigned as such is then incorporated into the likelihood function.[22] Since the empirical distribution has fat tails, with observed growth rates of absolute value greater than 10 log points, I allow for infinite support.

The other observations that do not follow the latent distribution are those that go missing in the next

[22]Missing and bunching observations could be assigned to any value of g^*. Identification uses the count of missing and the count of omitted and not the location of either.

year. I allow for 3 channels through which these organizations go unobserved. First, organizations do not file any information return if log receipts are below $\log(25,000)$. I drop the few observations with reported receipts below $rmin := \log(25,000)$ and set the share of truncated observations equal to the value taken by the latent conditional cdf at $rmin - r$.[23] Second, some share λ of current filers will not appear in the next year's data file regardless of their receipts, either because they miss the filing deadline or because their data is lost in some stage of the collection process. Third, I allow that an additional share δ go missing when crossing *notch*. In each case growth is unobserved, so for these observations I set the value of $g*$ equal to the minimum observable growth $(rmin - r)$.

Finally, I set $\theta = 0$ after obtaining nonparametric mode estimates between 0 and 0.005 for all years. The observed conditional cdf is therefore

$$F^*(g^*|r) = \begin{cases} 0 \\ \lambda + (1-\lambda) F(rmin - r|r) + \delta(1 - F(notch - r|r)) \\ \lambda + (1-\lambda) F(g^*|r) + \delta(1 - F(notch - r|r)) \\ \lambda + (1-\lambda) F(notch - Bwidth - r|r) + \delta(1 - F(notch - r|r)) \\ (1-\lambda)\left[F(notch - r + Rwidth|r) - F(notch - r - Bwidth|r)\right] \\ \qquad + b(1-\lambda-\delta)\left[F(notch - r + Rwidth|r) - F(notch - r|r)\right] \\ (1-b)(1-\lambda-\delta)\left[F(notch - r + Rwidth|r) - F(notch - r|r)\right] \\ \lambda + \delta + (1-\lambda-\delta)(F(g^*|r)) \end{cases}$$

[23]Results are robust to further truncation of the sample at $\log(100,000) - 1 \approx 37,000$, which would avoid any potential concerns about selective entry just above the truncation point. One could also exclude observations with current receipts in the omitted region or allow the density to be discontinuous in current receipts at r, in keeping with the potential concern that even the upper counterfactual region has been affected by repeated bunching at the notch, as discussed in Appendix A. In practice these adjustments also appear to have little effect on the estimates.

$$\begin{cases} for \quad g* < rmin - r \\ for \quad g* = rmin - r \\ for \quad rmin - r < g* < notch - r - Bwidth \\ for \quad notch - r - Bwidth \le g* < notch - r \\ for \quad g* = notch - r \\ \\ for \quad g* = notch - r + Mwidth \\ for \quad notch - r + Mwidth < g* \end{cases}$$

Maximizing the likelihood function $\sum_{i=1}^{N} \log\left[f^*\left(g_i^*|r_i\right)\right]$, where $f^*\left(g_i^*|r_i\right)$ is the discrete-continuous implementation of the conditional likelihood implied by $F^*\left(g_i^*|r_i\right)$, gives an estimate of the value of each parameter. For any value of r one can then obtain counterfactual growth estimates by plugging the desired value(s) of g into the estimated distribution function(s). Integrating over r gives the total counterfactual mass for the next year. I perform the estimation on observations with $r < 14 \approx notch + 2.5$ to reduce computation time and keep the results from being influenced too heavily by charities far above the notch. I rescale the resulting estimates of excess and reduced mass to represent shares of the full population in the next year (for comparison with static estimates).

Figure C.1 shows the fit of the model to the data for a sample of charities with log gross receipts below the notch by .24 to .25. Data for the omitted region has been dropped, and this is reflected in the MLE prediction. Otherwise the distribution is simply fit to the data as is.

Table 1: Comparison of Information Provided on IRS Forms for Charities

	Form 990-EZ	Form 990
Pages	3	9+
Revenues	15 lines	25 lines
Expenses	8 lines	5 lines
Statement of Functional Expenses		˜80 cells
Balance Sheets	8 lines	40 lines
Reconciliation with Audited Financials		if ∃ audited financials
Officers, Directors, Trustees, & Employees	Compensation	Compensation, # of relations
Compensated Former Officers, Directors, etc.		✔
Income Lines By Related vs. Unrelated		✔
Form 990-T if Unrelated Income > $1000	✔	✔
Controlled Entities		✔
Hours to Complete (Paperwork Reduction Act)	164	260

Table 2: Summary Statistics

All Public Charities, All Years (N=4,299,984)

	Lower Quartile	Median	Upper Quartile
Gross Receipts ($ Thousands)	72	200	825
Expenses ($ Thousands)	51	153	647
Assets ($ Thousands, Year-End Total)	36	180	996

Charities With Receipts of $80-130,000, FY2007 (N=36,173)

Major NTEE Category	**Share**	**Minor NTEE Category**	**Share**
Education	19.4%	Parent Teacher Group	6.5%
Arts, Culture, and Humanities	12.4%	Education - Single Organization Support	4.1%
Recreation, Sports, Leisure, Athletics	12.1%	Religion - Christian	3.9%
Human Services - Multipurpose and Other	9.6%	Baseball, Softball (Includes Little Leagues)	3.1%
Religion Related, Spiritual Development	7.8%	Fire Prevention/Protection/Control	1.9%
Community Improvement, Capacity Building	5.1%	Animal Protection and Welfare	1.8%
Housing, Shelter	4.1%	Education - Scholarships, Student Financial Aid, Awards	1.7%
Health	3.7%	Community/Neighborhood Development, Improvement	1.6%
Philanthropy, Voluntarism, Grantmaking Foundations	3.5%	Ameteur Sports Clubs, Leagues	1.3%
Public Safety	2.5%	Theater	1.2%
Animal-Related	2.4%	Soccer Clubs/Leagues	1.1%
Environmental Quality, Protection, and Beautification	2.3%	Community Service Clubs	1.1%

Table 3: Static Bunching Estimates: Distortions of the Income Distribution in the Pooled Sample

	Basic	Discontinuous	Two-Sided
Excess mass below the notch (*100)	.148***	.135***	.152
	(.020)	(.029)	(.093)
Bunching ratio (*100)	.592***	.537	.608*
	(.096)	(.123)	(.312)
Reduction in mass above the notch (*100)	.250***	.223**	-.055
	(.026)	(.049)	(.066)

Notes: The table shows deviations of the binned income distribution from a counterfactual estimated in the range of $50-200,000. In the Basic specification, the counterfactual is a cubic in gross receipts. The Discontinuous specification allows for a discontinuity at the notch, and the Two-Sided specification allows for a separate quadratic on each side of the notch. The excess mass shows the estimated extra share of charities with incomes below the notch relative to the counterfactual, the bunching ratio is the ratio of the excess mass to the counterfactual density at the notch, and the reduction above the notch is the difference between the counterfactual and actual share above. The Basic specification indicates that .148 percent of charities appear below the notch when they shouldn't, which is roughly equal to the number of charities that should be above the notch by up to $600 (=$100,000*.00592). The reduction in the number of charities above the notch is significantly larger than the addition below the notch, suggesting either misspecification or missing observations, and the flexible specifications do not reconcile the two results. The sample includes observations in years up to 2007 for charities also appear in the prior year (for comparability with the dynamic estimates). Bin width = $250. N = 969,842 in the range used for estimation and 2,907,476 total.

Table 4: Repeated Bunching: Charities Remain Just Below the Notch for Years

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
bunchbin	1.55***	1.15***	0.68***	0.63***	0.44*	0.20	0.31**	0.29**	0.11	0.14
	(0.37)	(0.30)	(0.24)	(0.21)	(0.24)	(0.20)	(0.15)	(0.14)	(0.12)	(0.14)

Notes: The table shows the results of regressing a dummy for remaining in the same log receipts bin (t) years in the future on a dummy for being in the bin just below the notch, with controls for year and a quadratic function of log receipts. The coefficients, which are multiplied by 100, show the heightened probability that charities just below the notch remain where they are. The sample includes charities within one log point of the notch in any starting year from 1990 to 1997. Standard errors are clustered by state. Bin width = .05. N=595,478.

Table 5: The Effect of Approaching the Notch on Organizational Finances

	(1)	(2)	(3)	(4)	(5)	(6)
	Receipts	Cross	IV: Receipts	Revenue	Expenses	Assets
Near Notch	-0.0017***	-0.0408***		0.0021	0.0016	-0.0009
	(0.0002)	(0.0033)		(0.0022)	(0.0032)	(0.0037)
crosslead1			0.0423***			
			(0.0031)			
N	1,076,302	1,076,302	1,076,302	1,070,904	1,069,204	1,064,645
Adj. R-Squared	0.999	0.001	1.000	0.383	0.078	0.037

Notes: The table shows the results of regressing financial variables on a dummy ("Near Notch") for bins that straddle the notch in future receipts, controlling for bins of growth rate (of width .1) each interacted with a quadratic function of current receipts. The negative relationships for growth of log receipts (1) and crossing the notch (2) reflect downward distortions of receipt growth in the neighborhood of the notch. Using the "Near Notch" dummy as an instrument for crossing (3) shows receipt growth is reduced by an average of .45 log points among charities induced not to cross. Effects on the growth of total revenue (4), expenses (5), and assets (6), all in logs, is not precisely estimated. The sample includes all charities growing by 0 to 1 log points. Standard errors are clustered by state.

Table 6: The Effect of Approaching the Notch on the Probability of Further Growth Years Ahead

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Near Notch	-0.053***	-0.021***	-0.018**	-0.015**	-0.016**	-0.017***	-0.016**	-0.015**	-0.017***	-0.022***	-0.012*	-0.014***
	(0.009)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.008)	(0.006)	(0.005)	(0.006)	(0.007)	(0.005)
N	307,526	260,209	261,771	256,548	252,669	247,364	245,228	240,193	234,728	231,303	225,570	221,296

Notes: The table shows the results of regressing a dummy for crossing the level of growth corresponding to the notch ("Cross" as defined in the text) (t) years in the future on the "Near Notch" dummy for bins that straddle the notch in the next year, controlling for bins of growth rate (of width .1) and a quadratic function of current receipts. The coefficients show charities a signficant reduction of at least one percentage point in the probability of crossing the notch at all horizons. The sample includes charities within one log point of the notch in any starting year from 1990 to 1997 and growing by 0 to 1 log points. Standard errors are clustered by state.

Table 7: Heterogeneity in Share Crossing the Notch, by Size

	(1)	(2)	(3)	(4)	(5)	(6)
Near Notch	-.260***	-.197***	-.178***	-.371***		
	(0.073)	(0.042)	(.024)	(.077)		
Log Total Revenue * Near Notch	.020***			-.0001		.082
	(0.007)			(.012)		(.150)
Log Total Revenue	-.002			.005**		.005***
	(0.0101)			(.002)		(.002)
Log Expenses * Near Notch		.014***		.018***		.206
		(0.004)		(.007)		(.151)
Log Expenses		-.007***		-.009***		-.009***
		(0.002)		(.001)		(.001)
Log Assets * Near Notch			.013***	.013***		.033
			(0.002)	(.002)		(.033)
Log Assets			-.005***	-.005***		-.005***
			(0.001)	(.001)		(.001)
990-EZ * Near Notch					-.129***	-3.619***
					(0.042)	(1.320)
990-EZ					.008	.009
					(0.012)	(0.060)
N	1,071,602	1,070,546	1,068,105	1,059,710	1,053,004	1,036,868

Notes: The table shows the results of regressing a dummy for crossing the level of growth corresponding to the notch ("Cross" as defined in the text) on a dummy ("Near Notch") for bins that straddle the notch in future receipts, interacted with various measures of size, and controlling for bins of growth rate (of width .1), each interacted with a quadratic function of current receipts. The positive coefficients on the interaction terms indicate that larger charities are less likely to reduce income to stay below the notch when first approaching it. Columns (5) and (6) report results of regressions that only include observations moving to the notch if they appear in the IRS Statistics of Income 990-EZ sample, thereby excluding those already filing Form 990. The restriction renders the interaction terms insignificant but has little effect on point estimates. The sample for all regressions includes charities growing by 0 to 1 log points. Standard errors clustered by state.

Table 8: Heterogeneity in Share Crossing the Notch in the Short Run, by Staffing

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Has "Compensation" * Near Notch	-0.0257***	-0.0247**		-0.0090		-0.0015	-0.0035
	(0.0079)	(0.0115)		(0.0112)		(0.0114)	(0.0115)
Has "Compensation"	0.0132***	0.0089***		0.0113***		0.0117***	0.0115***
	(0.0013)	(0.0024)		(0.0024)		(0.0025)	(0.0026)
Has "Other Salary" * Near Notch			-0.0509***	-0.0483***			-0.0366**
			(0.0118)	(0.0117)			(0.0174)
Has "Other Salary"			-0.0043**	-0.0076***			-0.0071**
			(0.0019)	(0.0019)			(0.0027)
Has Payroll Tax * Near Notch					-0.0468***	-0.0460***	-0.0185
					(0.0101)	(0.0097)	(0.0152)
Has Payroll Tax					-0.0007	-0.0058***	-0.0007
					(0.0019)	(0.0019)	(0.0028)
N	989,706	355,810	355,810	355,810	355,810	355,810	355,810
Adj. R-Squared	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Notes: The table shows the results of regressing a dummy for crossing the level of growth corresponding to the notch ("Cross" as defined in the text) on a dummy ("Near Notch") for bins that straddle the notch in future receipts, interacted with dummies for different types of staffing. Staffing is only known for filers of Form-990 and is defined for each charity in its first year with receipts above the notch. The negative coefficients on the interaction terms indicate that charities with administrative staff are less likely to cross the notch when first approaching it. Controls include dummies for bins of growth rate (of width .1) each interacted with a quadratic function of current receipts. The sample includes all charities with current growth between 0 to 1 log points that ever appear above the notch. Regressions (2) through (7) include only charities that first appear above the notch in or after 1997, the year in which "Other Salary" and "Payroll Tax" first appear in the data. Standard errors are clustered by state.

Table 9: MLE Estimates of Propensities to Manipulate Income Or Leave the Sample

	Dynamic (1)	Dynamic (2)	Dynamic (3)	Static
Share bunching from below notch	0.087***	0.090***	0.079***	
	(0.003)	(0.003)	(0.005)	
Share bunching from above notch	0.001	0.000	0.000	
	(0.002)	(0.000)	(0.000)	
Extra share from below going missing		0.013***	0.009***	
		(0.001)	(0.002)	
Extra share missing instead of crossing			0.017***	
			(0.004)	
Excess mass below the notch (*100)	.183***	.184***	.159***	.148***
	(.030)	(.026)	(.010)	(.020)
Additional reduction above the notch (*100)	.183***	.184***	.246***	.250***
	(.030)	(.026)	(.018)	(.026)
AIC	5,531,280	5,530,760	2,765,374	

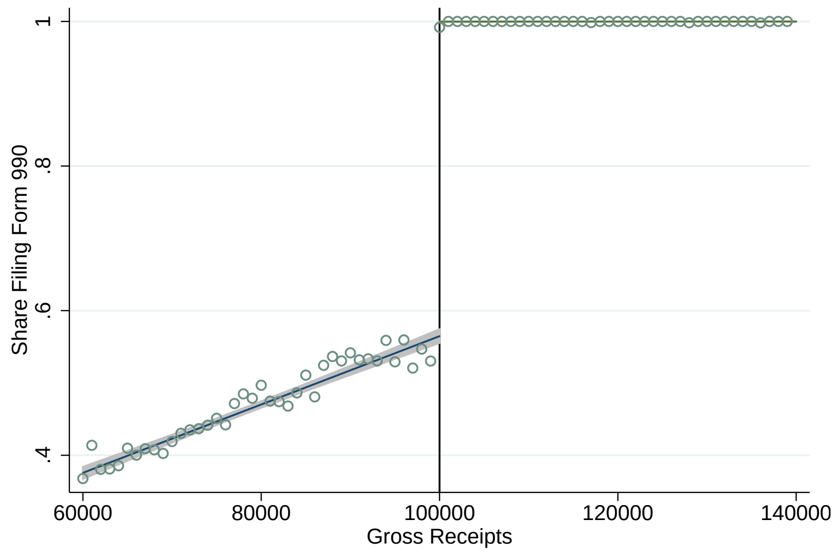
Notes: The table shows the results of maximum likelihood estimation of the bunching propensities of charities approaching the notch next year from current receipts below or above the notch, with and without adjusting the likelihood function for an increased share of charities with current receipts below the notch leaving the data (either in total or in proportion to how frequently they should cross over the notch). The top two parameter estimates indicate that charities that approach the notch from below are significantly more likely to manipulate receipts to remain below the notch in the next year. Allowing more attrition among observations below the notch slightly reduces the bunching propensity estimate and lowers the excess mass in the bunching region in the next year. The estimates in columns (2) and (3) require the excess mass below the notch to equal the reduction above the notch and hence obtain estimates in between the static estimates of these two values, whereas the dynamic estimates allowing for extensive-margin responses of those who would have crossed gives similar results to the static estimate. All regressions allow for attrition that is linear in current receipts and manipulation of receipts in the range $80-130,000. Standard errors calculated using the Delta Method. N=2,907,476.

Table 10: Robustness of Bunching Estimates To Omitted Range

	Dynamic	Static
Reduced Range: $100-$110K	.111***	.208***
	(.007)	(.018)
Reduced Range: $100-$120K	.151***	.169***
	(.011)	(.019)
Reduced Range: $100-$130K	.159***	.148***
	(.010)	(.020)
Reduced Range: $100-$140K	.170***	.136***
	(.010)	(.022)
Reduced Range: $100-$150K	.164***	.146***
	(.011)	(.024)

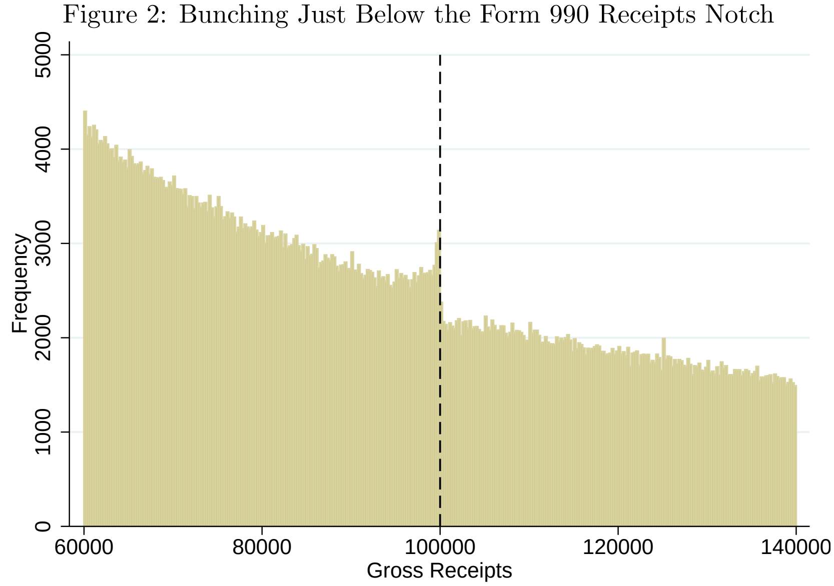
Notes: The table shows the results of maximum likelihood dynamic estimation and static estimation of (100 times) the excess mass in the $80-100,000 bunching region for various widths of the reduced region from which this mass has moved. According to the dynamic approach, the number of extra charities in the bunching region is about .2 percent of all charities in the sample. The dynamic estimates have the expected pattern: bunching is underestimated when the reduced range from which charities bunch is not sufficiently large, but estimates are stable as this range is widened. Static estimates are large when the specified reduced region is too small because the counterfactual is underestimated but also stabilize when the range is widened. For comparability, static estimation is performed on the sample of organizations that appeared in the previous year. The reduced range is $100-130,000 for all regressions. Standard errors calculated using the Delta Method. N=2,907,476.

Figure 1: Probability of Filing Form 990 Around the Receipts Notch



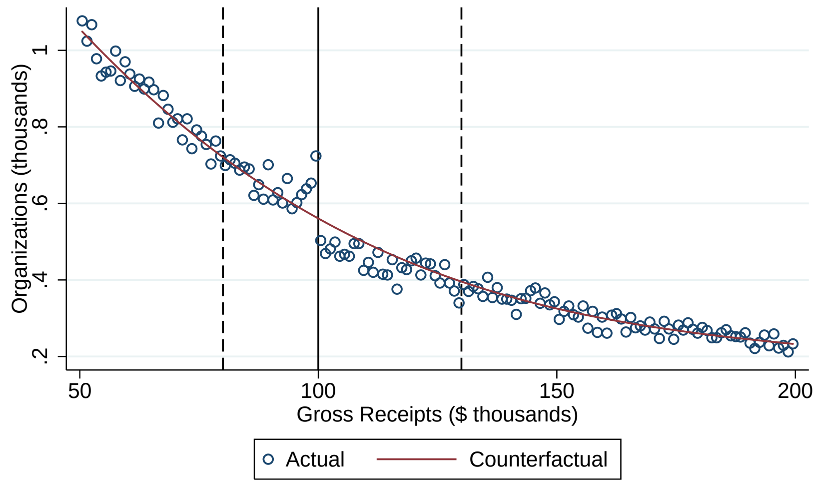
Notes: The figure shows the results of regressing a dummy for those filing Form 990 (vs. 990-EZ) in 2007 on quadratics in gross receipts below and above the $100,000 notch at which charities lose eligibility to instead file Form 990-EZ. Curves with standard error bands show the results of these regressions and circles show the mean within a $1000 receipts bin. The share of organizations filing the longer form is increasing in receipts up to the notch, with nearly 100% compliance above the notch. N=72,354.

Figure 2: Bunching Just Below the Form 990 Receipts Notch



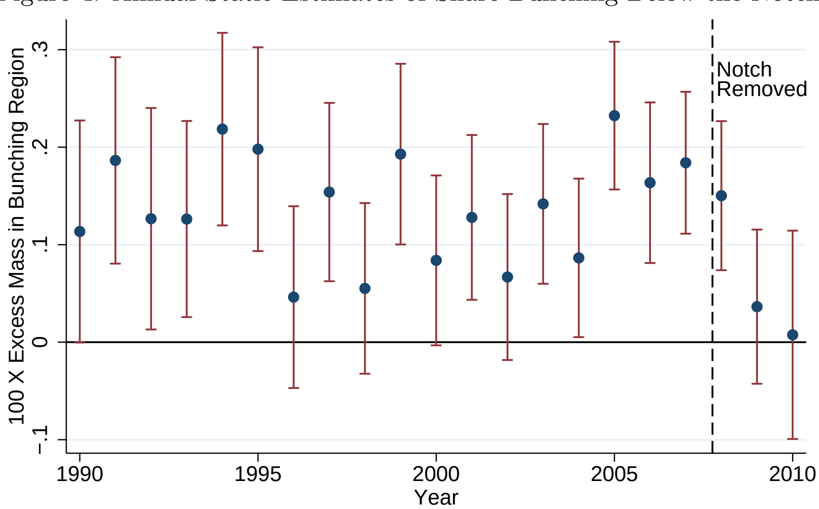
Notes: The figure is a histogram of gross receipts. An excess of charities just below the $100,000 notch appears as bunching in what is otherwise a smooth distribution. N=810,869. Bin width=$250. Years 1999-2007 pooled.

Figure 3: Distribution of Receipts in 2006 vs. Smooth Counterfactual



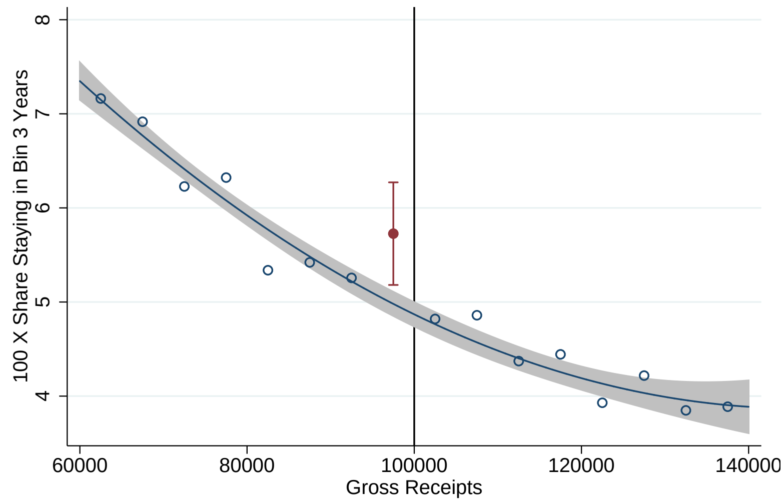
Notes: The figure shows the deviation of the 2006 distribution, represented by a histogram in blue circles, from a smooth counterfactual. Each bin is treated as an observation. Bin counts are regressed on a polynomial of degree 3, which estimates the counterfactual distribution, and a dummy variable for each bin in the omitted range of $80-130,000 indicated by the dashed lines. Excess "bunching" mass is calculated as the sum of coefficients on dummy variables for each bin in the bunching region between the dashed line at $80,000 and the solid at the the $100,000 notch. Similarly, the estimated reduction in mass above the notch is the sum of coefficients on dummies for each bin up to $130,000. N(graph)=92,791. N(2006)=264,770. Bin width=$1000.

Figure 4: Annual Static Estimates of Share Bunching Below the Notch



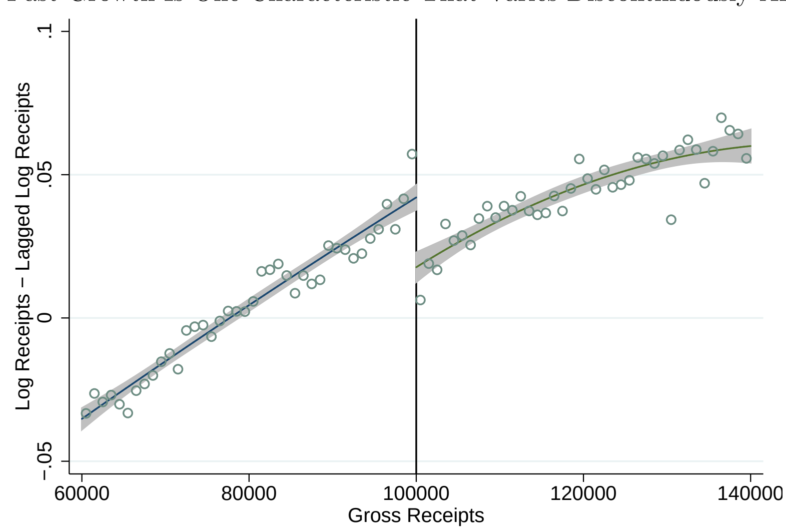
Notes: The figure shows the excess mass of charities below the notch in each year, as estimated using the static approach described in the text and Figure 3. Circles indicate the estimates and lines show 95-percent confidence intervals. Estimates fluctuate somewhat around the pooled estimate of .148. There is significant bunching below $100,000 for one year after the notch was raised to $1,000,000 (where new bunching forms), suggesting slow adjustment or lack of understanding that the notch had moved. The counterfactual for each year is a polynomial of degree 3 estimated on observations with receipts of $50-200,000 but outside of an omitted region of $90-130,000. Year is the calendar year in which the charities' fiscal years begin. The sample consists of charities that appear in the prior year (for comparability to other estimates in the paper). N=2,907,476.

Figure 5: Repeated Bunching: Share Staying Within Bin For 3 Years



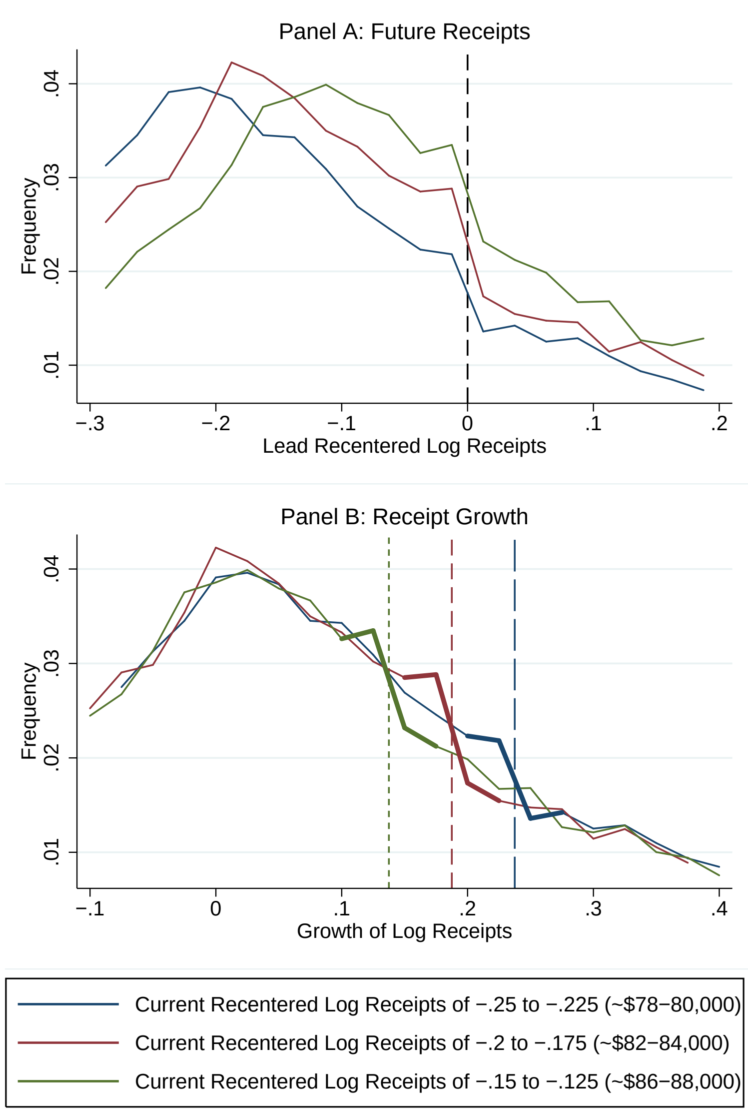
Notes: The figure shows the results of regressing a dummy for remaining in the same $5000 receipts bin 3 years after the current year on a quadratic in gross receipts and a dummy for the bin just below the notch. The marker with a 95-percent confidence interval shows that organizations in the bunching region just below the notch are especially likely to remain where they are for several years. Standard errors clustered by state. N=329,448. Bin width=$5000.

Figure 6: Mean Past Growth Is One Characteristic That Varies Discontinuously Around the Notch



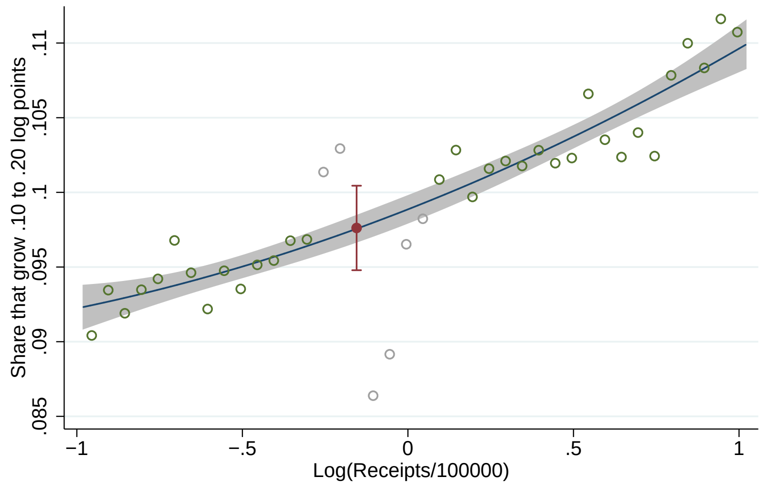
Notes: The figure shows the results of regressing growth of log receipts (from the previous year to the current year) on quadratics in gross receipts below and above the $100,000 notch. Curves with standard error bands show the results of these regressions and circles show the mean within a $1000 receipts bin. Mean growth is a discontinuous function of current receipts, so traits of charities that correlate with past income may also appear distorted in current income. This figure motivates conditioning on past income when describing and measuring income manipulation around the notch. N=688,948.

Figure 7: Distorted and Undistorted Sections of Conditional Distributions of Future Receipts/Growth



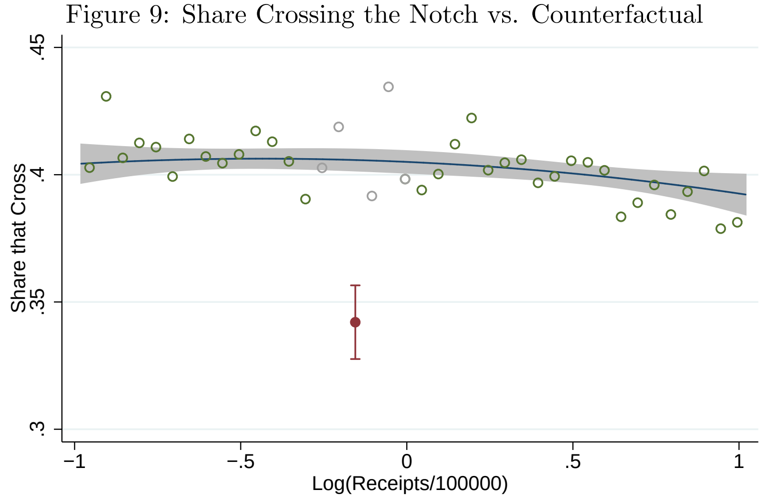
Notes: The figure shows the distribution of future receipts (Panel A) and growth to future receipts (Panel B) for charities in three sample bins of current receipts. The distributions for each group exhibit a spike at incomes just below the notch and a depression just above it, indicating manipulation of future income in order to stay below the notch. The growth distribution of each group is similar except around the notch, which appears in a different part of each distribution. Because the growth distribution does not vary too much with current income, the extent of distortion in the rates of growth that bring charities with one level of current receipts to the notch can be identified using the likelihood of such growth rates among charities with a different level of current receipts. N=92,242. Bin width = .025.

Figure 8: Share Growing To a Range That Spans the Notch is Unaffected



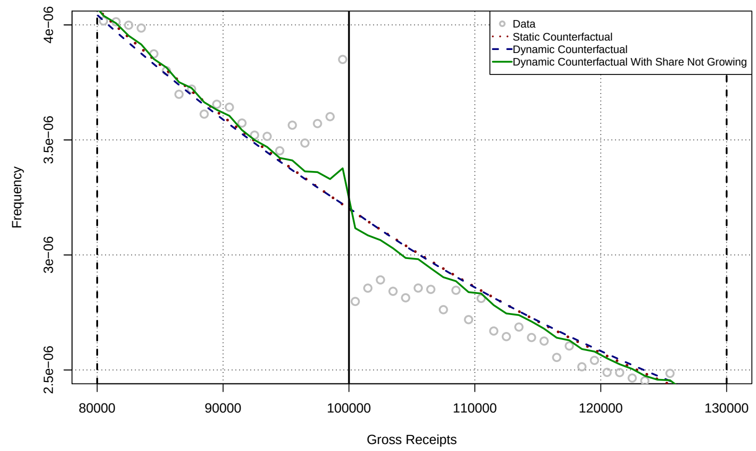
Notes: The figure shows the results of regressing the probability of growing log receipts by .1 to .2 (from the current year to the next) on a quadratic in current recentered log receipts and a dummy ("Near Notch") for the bin for which future receipts lie in the "omitted range" straddling the notch. The marker with a confidence interval represents the average among the "Near Notch" bin. Because growth of .1 to .2 log points from this bin leads to receipts on both sides of the notch it includes both those who manipulate and those who don't and so the overall probability of growth in this range is unaffected. Charities in the "Near Notch" bin can therefore be compared to counterfactuals constructed using charities in the same growth range but with higher and lower current receipts. Comparisons should exclude charities in bins represented by light markers because manipulation of income from one side of the notch to the other alters the sample with growth of .1 to .2 from these bins. The same arguments apply to other growth ranges. N=152,191. Omitted range is -.08 to .07. Bin width = .05.

Figure 9: Share Crossing the Notch vs. Counterfactual



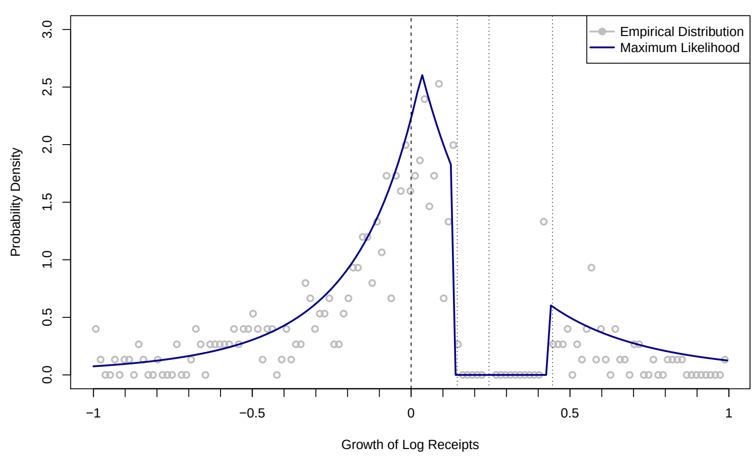
Notes: The figure shows the results of regressing the probability of crossing the growth rate corresponding to the notch (the "Cross" dummy described in the text) on a quadratic in current recentered log receipts and a dummy ("Near Notch") for the bin for which future receipts lie in the "omitted range" straddling the notch. The marker with confidence intervals represents the average among the "Near Notch" bin. Charities that move to a range near the notch reduce their income to stay below it and are therefore less likely to cross it than predicted by the estimated counterfactual represented by the curve. The figure sample consists of organizations growing .1 to .2 log points, and the same result obtains for other ranges of positive growth. N=152,191. Omitted range is -.08 to .07. Bin width = .05.

Figure A.1: Smooth and Non-smooth Counterfactual Distributions of Income



Notes: The figure shows the distribution of gross receipts and some potential counterfactual distributions. The dotted and dashed lines show the estimated counterfactual using the static and dynamic approaches, which give similar results. The solid line shows a counterfactual in which 80 percent of charities grow according to the conditional distribution estimated by maximum likelihood and 20 percent have no growth. When a share of charities don't grow the counterfactual is not smooth around the notch, implying different estimates and interpretation of the excess mass observed in the data. Details of the dynamic estimates are provided in Appendix C. The plot includes observations that appear in a prior year, the static counterfactual is estimated directly from this data, and the dynamic counterfactuals apply the estimated distribution of growth conditional current income to the distribution of incomes in the prior year.

Figure C.1: Estimation of the Distribution of Growth Rates



Notes: The figure shows the distribution of growth in log receipts. The curve shows the fit of the maximum likelihood estimate of this distribution to the data represented by the histogram in circular markers. The sample consists of organizations .24 to .25 log points below the notch, implying that growth of about .245 puts these charities near the notch in the next year, as represented by the middle dashed vertical line. Observations in the omitted region around the notch, marked by the surrounding dashed lines, have been omitted from estimation of the shape. These observations are only used to compare the number of organizations just above and below the notch to the numbers implied by the counterfactual to get estimates of the share bunching or missing. N=12,637.