The Cost of Requiring Charities to Report Financial Information

Marx, Benjamin M.

University of Illinois at Urbana-Champaign

25 August 2018

The Cost of Requiring Charities to Report Financial Information

Benjamin M. Marx*

July 2018

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 provides a theoretical model for evaluating welfare effects of moving such thresholds. It then analyzes an income notch at which IRS reporting requirements for charitable organizations become more onerous. Standard bunching estimates imply that the average charity will reduce reported income by $750 to $1000 to avoid filing the more onerous information return. Panel data methods show that an even larger share of charities fail to appear 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, suggesting the report imposes an adjustment cost on new filers.

JEL: L38, H26, D64.

* Department of Economics, University of Illinois at Urbana-Champaign, 214 David Kinley Hall, 1407 W. Gregory, Urbana, IL 61801. Correspondence should be sent to benmarx@illinois.edu. I am especially grateful to Bruce Kogut, Wojciech Kopczuk, Robert McMillan, Brendan O'Flaherty, Bernard Salanié, and Miguel Urquiola for invaluable advice and discussions. I also thank Raj Chetty, Julie Berry Cullen, Jonathan Dingel, Francois Gerard, Jessie Handbury, Todd Kumler, Ilyana Kuziemko, Corinne Low, David Munroe, Giovanni Paci, Abigail Payne, Petra Persson, Jonah Rockoff, Maya Rossin-Slater, Ugo Troiano, Lesley Turner, Eric Verhoogen, Reed Walker, Caroline Weber, and seminar participants at meetings of the International Institute for Public Finance, the Office of the U.S. Department of the Treasury Office of Tax Analysis, the National Tax Association Annual Conference, the Association for Public Economic Theory, the Midwestern Economics Association, and the University of Illinois and other universities. I gratefully acknowledge the financial support of a Nonprofit Research Fellowship from the National Bureau of Economic Research and the Outstanding Dissertation Award from the National Tax Association.

1 Introduction

Taxes, price discrimination, income eligibility limits, and other policies create discontinuities in budget constraints, or "notches" (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 "bunching" below the notch by reducing (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), the Affordable Care Act, value-added taxes (Keen and Mintz (2004), Dharmapala et al. (2011)), and some countries' payroll tax systems (Dixon et al., 2004). Income notches produce deadweight loss if they affect income, and a growing literature estimates behavioral parameters by quantifying bunching in the distribution of income around a notch.

This paper studies bunching by charitable organizations at an income eligibility limit for simplified IRS reporting forms. The charitable sector is large, accounting for about 9 percent of all U.S. wages and salaries (Roeger et al., 2012). Charities receive exemption from paying income taxes because they provide public goods. However, they must file annual information returns with the IRS to enable monitoring by the government and the public. Some monitoring is likely warranted to promote good management and to prevent "for-profits in disguise" from exploiting the organizational form to evade taxes. However, requiring charities to report this information imposes a compliance cost. The problem of determining optimal reporting requirements for these organizations mirrors regulations in other industries, particularly where firms likely produce sizable externalities.

I first provide a theoretical framework to describe welfare effects of changing information reporting requirements. Organizations maximize net available resources, making the model applicable to charities promoting their missions or other types of organizations. Those exceeding an income threshold must pay a fixed cost to report more information than those below the threshold. The government sets the location of the threshold so as to maximize a social welfare function that aggregates the organizations' net resources and the value of the information they provide. The first-order conditions from the government's problem produce sufficient statistics for the welfare effect of moving the policy threshold. I begin with a static model, which provides the intuition that requiring more information from organizations provides a public information good but increases the compliance cost imposed on marginal organizations. I then general the model to two periods and then to the infinite horizon and allow for extensive-margin responses. This identifies a more general set of parameters to be estimated.

A variety of empirical tools provide information about the effects of the reporting notch on charities.

For completeness, I include results of standard bunching estimation techniques. Marx (2018) shows that the standard approach is subject to considerable bias in the presence of extensive-margin responses or a notch that is faced repeatedly. Marx (2018) proposes three new bunching designs that exploit panel data, and I employ all of them here. The first takes advantage of a one-time notch, which charities faced at particular income thresholds in 2008 and 2009 as the notch was moved, to estimate long-run effects of approaching a notch. The second estimates an OLS counterfactual through bins of agents' income from two years, rather than one, to describe heterogeneity in responses to the notch and the extent to which agents are induced to bunch repeatedly. This design has already been employed by St.Clair (2016). The third design employs the same pairs-of-years identification strategy as the second design but uses maximum likelihood estimation to precisely quantify the bunching response and extensive-margin responses (in this setting non-filing or late filing). Last, I employ donut-RD specifications, which exclude the selected sample in the omitted range surrounding the notch, to estimate the average number of years that it takes for extensive-margin responders to return to the sample.

Results from each of the empirical designs are illuminating. I show that the reporting notch reduces the incomes of public charities in both the short and the long run. I find that bunching of charities at this notch permanently reduces income, indicating that charities actually forgo income to bunch (and are not simply misreporting). There is significant heterogeneity in the response, with evidence that much of the compliance cost is a one-time adjustment; 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.. Extensive-margin responses appear to be at least as important as the bunching response. Extensive-margin responders leave the sample for an average of 1.5, again suggesting a temporary adjustment to filing the more comprehensive information return. Welfare analysis indicates that the income threshold for filing should be lowered if the social benefit of each report filed exceeds \$55. With standard methods one would vastly overestimate this quantity to be over \$750.

The findings of this study provide new information about the behavior of charities and effects of regulating the sector. Fack and Landais (2016b) show that charitable donations of those at the top of the income distribution fell when the Tax Reform Act of 1969 introduced new regulations for charitable foundations. Marx (2015) finds the same result in the receipts of foundations and provides evidence that this decline involved both the desired deterrence of non-charitable activity (as Fack and Landais (2016a) conclude for a French tax reform requiring receipts for donations) and some undesired reductions caused by imposing compliance costs. Foundations face different rules than the public charities studied here, but the regulatory trade-offs are similar. Evidence of the compliance costs faced by public charities is provided by Blumenthal and Kalambokidis (2006a) through a survey of charities and by St.Clair (2016) through estimation of bunching at a state audit threshold. Both find compliance costs to be non-negligible, consistent with my results for national reporting requirements.

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, 2006b). 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. Moreover, regulatory instruments may affect the firm size distribution and its evolution. 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)).

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 provides background information about the information reporting of public charities and the panel of data this provides. In Section 3, I model the welfare effects of changing the income threshold at which a regulation applies. Section 4 presents empirical analysis and welfare implications. Concluding remarks appear in Section 5.

2 The Setting: Nonprofit Information Returns

This section describes the context of a reporting notch for U.S. charities and the data from these organizations. 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

Recent Congressional hearings and increases in IRS monitoring of the nonprofit sector demonstrate renewed interest in the optimal regulation of charitable organizations.[1] Public charities are organizations granted income and sales tax exemption under section 501(c)(3) of the Internal Revenue Code on condition that they serve a public purpose and do not distribute profits.[2] 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. 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.[3]

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 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 (2006b) asked for

[1] 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, revoking the tax-exempt status of more than a quarter-million organizations that had not filed in the three years leading up to 2011, and the introduction in 2014 of a new form 1023-EZ to simplify the process for small charities to apply tax-exempt status. 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 for charitable hospitals.

[2] 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.

[3] 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.

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.[4] 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 (2006b) 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

[4]Kline and Tartari (2015) do not find bunching at the earnings eligibility threshold for a state welfare reform despite other evidence of effects on earnings. They argue bunching may understate total responses for individuals who have incomplete control over earnings and whose incentives to reduce taxes by underreporting may result in reported earnings just above the threshold. While charities undoubtedly face some lumpiness in each of their sources of income, they are likely to have significantly more control over earnings than an individuals negotiating with employers, and since revenues are not taxed there is no obvious reason to underreport except to keep income below reporting thresholds.

[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 (2015) 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 percent of firms in this region must file the longer form because their assets are above the $250,000 notch. The others 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 files 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 (2013).[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 percent of charities with assets between $200,000 and $250,000 have receipts over $100,000, and an additional 15 percent in this range file the full Form 990 by choice. I therefore focus on the receipts notch in the model and empirical analysis.

3 Theoretical Framework

To more fully evaluate the welfare effects of adjusting a notch policy I first extend the model to the infinite horizon and incorporate extensive-margin responses. I then incorporate the parameter estimates for the charity reporting notch to quantify the marginal social cost of filing and hence the minimum marginal social

[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.

benefit for which increasing the number of filers would enhance welfare.

This section presents a conceptual framework for using bunching estimates for welfare analysis in static and dynamic settings. The model can be applied generally to income eligibility limits that impose a cost (that is possibly unobserved by the government and econometrician) but is described as a charity reporting notch for concreteness. Policy design weighs the social value of obtaining information through reporting requirements against the costs these requirements impose, including the avoidance costs of organizations that bunch. A static model provides intuition and illustrates the connection between regulatory and tax notches, and a two-period model reveals how dynamic considerations affect the choices of forward-looking agents. The optimal location of the notch is found to depend 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.

Static Model

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

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

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

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

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.

[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).

[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.

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 $\bar{\delta} = \max\left(\delta\left(\rho, \phi, \gamma, \bar{y}\right)\right)$. 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 $C_a\left(y, 0, \gamma, \bar{y}\right) \geq M$. Denote the excess mass observed below the notch as $B := \int \left[\int_{\rho}^{\rho+\delta} f\left(\bar{y}\right) d\bar{y}\right] g\left(\gamma\right) d\gamma = b\left(F\left(\rho+\bar{\delta}\right) - F\left(\rho\right)\right)$, where b is the share of organizations in the reduced region that choose to bunch. Assume that $\frac{db}{d\rho} \approx 0$, $\frac{d\bar{\delta}}{d\rho} \approx 0$, and $\exists\bar{\pi} : \pi\left(\phi, \gamma, \rho+\delta\right) \approx \bar{\pi} \forall \gamma, \bar{y} \in \left[\rho, \rho+\delta\right]$. In words, slight movements of the notch have little effect on the share of organizations that bunch or the maximum amount by which they will reduce income, and all reports from organizations have roughly the same social value. 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 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 $-\bar{\pi}\left[bf\left(\rho+\bar{\delta}\right) + \left(1-b\right) f\left(\rho\right)\right]$.[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

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

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 may vary across organizations (as reflected in the fact that only larger organizations, for which compliance costs are likely to be lower, are required to file the long form). 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 $b\left(F\left(\rho+\bar{\delta}\right) - F\left(\rho\right)\right) = B \approx b\bar{\delta}f\left(\rho\right)$ to back out an estimate of $\bar{\delta}$ from estimates of the counterfactual distribution and excess mass. Kleven and Waseem (2013) estimate a

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

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 the density is not very steep at the notch, I have followed the practice in other bunching studies and 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).

It is far more difficult to estimate the marginal avoidance 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 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 agents' preferences. 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.

Two Period Model

Individuals and firms often face the same notch repeatedly. Extending the model to two periods reveals that the behavior of forward-looking agents will generally differ from the static version of the problem, muddling the interpretation of static bunching estimates.

Suppose an agent lives for two periods and in each period faces the problem described above. Let $f\left(\bar{y}_2|\bar{y}_1, y_1\right)$, the density of optimal income in the second period, depend on optimal or chosen income from the first period. For example, if a charity contacts more potential donors in a given year then donations for both that year and the following year are likely to increase. The solution to the problem in period 2 will have the form shown above. Let $V_2\left(\rho, \phi, \bar{y}_2\right)$ denote the value of the objective achieved in year 2. The problem in the first year will have a similar solution, with charities bunching from just above the notch. The problem in the first year can be written as

$$V_1(\rho,\phi)=\max_{y_1}\left\{\beta EV_2(\rho,\phi,\bar{y}_2)+\begin{cases}y_1-C(y_1,0) & y_1\le\rho\\ y_1-C(y_1,y_1-\rho) & \rho<y_1\le\rho+\delta\\ y_1-C(y_1,0)-\phi & y_1>\rho+\delta\end{cases}\right\}$$

The first-order condition is now

$$0=\beta\frac{d}{dy_1}EV_2(\rho,\phi,\bar{y}_2)+\begin{cases}1-C_y(y_1,0) & y_1\le\rho\\ 1-C_y(y_1,y_1-\rho)-C_a(y_1,y_1-\rho) & \rho<y_1\le\rho+\delta\\ 1-C_y(y_1,0) & y_1>\rho+\delta\end{cases}$$

If $\frac{d}{dy_1}EV_2(\rho,\phi,\bar{y}_2)=0$ then the solution is exactly as in the static setting. In general this will not be the case if $\frac{d}{dy_1}f(\bar{y}_2|\bar{y}_1,y_1)\neq 0$, because $EV_2(\rho,\phi,\bar{y}_2)$ is increasing in $\bar{y}_2$. If we assume that $\frac{d}{dy_1}EV_2(\rho,\phi,\bar{y}_2)>0$ then the optimal value of y_1 will be greater than in the static case because it will be worth losing money on the last marginal dollar raised in period 1 in order to increase expected earnings in year 2.

Importantly for research, facing a notch more than once will also generally affect the amount of bunching. For example, consider a charity that only lives for one year and has potential income $\bar{y}=\rho+\delta$, making it indifferent between bunching and not bunching. If bunching is not achieved entirely by avoidance than this charity earns a higher income if it doesn't bunch than if it does bunch. If the charity were to then learn that it would live for another period and that income in this second period would be increasing in income earned in the first period then bunching would now be strictly dominated. In this case, δ becomes smaller and bunching decreases. If agents know or believe that income is persistent then estimates of bunching cannot be interpreted as the solution to a static problem, and this may explain small estimates in the literature.

Infinite Horizon With Extensive-Margin Responses

Now I expand the model to enable welfare calculations for the empirical setting. Charities are able to exist in perpetuity and to temporarily leave the sample to avoid compliance costs. To simplify exposition slightly I describe the case of homogeneous bunching ($b=1$). Let $f(y)=\int f(\bar{y}_t|\bar{y}_{t-1},y_{t-1})\,g(\gamma)\,d\gamma$ and $F(y)=\int F(\bar{y}_t|\bar{y}_{t-1},y_{t-1})\,g(\gamma)\,d\gamma$ denote the aggregate distribution of incomes.

Suppose the cost of going missing from the sample for a year (net of not filing, if charities are actually not filing versus filing late) is $\mu_i\in(-\infty,\infty)$, and this cost varies across charities with cumulative distribution $H(\mu)$. It will be assumed, based on the empirical results, that this cost is only negative for charities with $y_{t-1}<\rho$, i.e. that charities already appearing above the notch will not find it worthwhile to go missing. The fact that the distribution $H(\mu)$ will therefore depend on y_{t-1} will be suppressed for notational convenience. Two implications follow directly. First, if a charity grows to an income level above its reduced range then it will go missing with probability $H(0)$. Second, if we denote $\triangle := -1\left[\hat{y} - C(\hat{y}, \hat{y} - \rho) - (\bar{y} - C(\bar{y}, 0) - \phi)\right]$ then the probability of going missing when growing to the reduced range $(\rho, \rho + \delta)$ is $H(\triangle)$.

The dynamic-programming formulation of the charity's problem (with discount factor β) is

$$V_0(\rho, \phi, \gamma) = \max_{y,a} \left\{ y_0 - C(y_0, a_0, \omega) - \phi(\gamma, \omega) \cdot 1\{y_0 - a_0 > \rho\} + \sum_{t=1}^{\infty} \beta^t EV(\rho, \phi, \gamma) \right\}$$

$$= \max_{y} \{1\{\bar{y}_0 < \rho\}(\bar{y}_0 - C(\bar{y}_0, 0, \gamma, \bar{y}_0)) + 1\{\bar{y}_0 \geq \rho + \delta\}(\bar{y}_0 - C(\bar{y}_0, 0, \gamma, \bar{y}_0) - \phi(\gamma, \bar{y}_0) - 1\{\mu < 0\}\mu)$$

$$+1\{\rho \leq \bar{y}_0 < \rho + \delta\}\left[1\{\mu \geq \triangle\}(\hat{y}_0 - C(\hat{y}_0, \hat{y}_0 - \rho, \gamma, \bar{y}_0)) + 1\{\mu < \triangle\}(\bar{y}_0 - C(\bar{y}_0, 0, \gamma, \bar{y}_0) - \phi(\gamma, \bar{y}_0) - \mu)\right]$$

$$+\sum_{t=1}^{\infty} \beta^t \int_0^{\rho} (\bar{y}_t - C(\bar{y}_t, 0, \gamma, \bar{y}_t)) f(\bar{y}_t) d\bar{y}_t + \sum_{t=1}^{\infty} \beta^t \int_{\rho+\delta}^{\infty} (\bar{y}_t - C(\bar{y}_t, 0, \gamma, \bar{y}_t) - \phi(\gamma, \bar{y}_t) - 1\{\mu < 0\}\mu) f(\bar{y}_t | \bar{y}_{t-1}, y_{t-1}) d\bar{y}_t$$

$$+\sum_{t=1}^{\infty} \beta^t \int_{\rho}^{\rho+\delta} \left[1\{\mu \geq \triangle\}(\hat{y}_t - C(\hat{y}_t, \hat{y}_t - \rho, \gamma, \bar{y}_t)) + 1\{\mu < \triangle\}(\bar{y}_t - C(\bar{y}_t, 0, \gamma, \bar{y}_t) - \phi(\gamma, \bar{y}_t) - \mu)\right] f(\bar{y}_t | \bar{y}_{t-1}, y_{t-1}) d\bar{y}_t \Bigg\}$$

As in the static model, raising the level of the notch has a first-order effect on utility when bunching:

$$\frac{d}{d\rho} V(\rho, \phi, \gamma) = 1\{\rho \leq \bar{y}_0 < \rho + \delta\} 1\{\mu \geq \triangle\}(C_a(\hat{y}_0, \hat{y}_0 - \rho, \gamma, \bar{y}_0)) + \frac{d}{d\rho} \sum_{t=1}^{\infty} \beta^t EV(\rho, \phi_t, \gamma)$$

$$= 1\{\mu \geq \triangle\}\left[C_a(\hat{y}_0, \hat{y}_0 - \rho, \gamma, \bar{y}_0) + \sum_{t=1}^{\infty} \beta^t \int_{\rho}^{\rho+\delta} \left[(C_a(\hat{y}_t, \hat{y}_t - \rho, \gamma, \bar{y}_t))\right] f(\bar{y}_t | \bar{y}_{t-1}, y_{t-1}) d\bar{y}_t\right]$$

The cost of raising the level of the notch is again the lost social benefit from forms no longer filed. Let $\prod(\rho) = 1\{\mu \geq 0\}\pi \sum_{t=1}^{\infty} \left[1\{\bar{y}_0 \geq \rho + \delta\} + \beta^t \int_{\rho+\delta}^{\infty} \pi f(\bar{y}_t | \bar{y}_{t-1}, y_{t-1}) d\bar{y}_t\right]$ be the present value of information benefits, where π is now assumed to be constant for simplicity. Also, define parameter $\eta_h = F(\rho + \delta | y_t = \rho)_{t+h} - F(\rho + \delta | y_t = \rho + \delta)_{t+h}$, the effect of bunching today on the probability of bunching h years in the future. Assume for simplicity that $\forall h, \eta_h = \eta$, i.e. that this effect is constant over horizon, consistent with the empirical results for this setting.

There are now two discrete responses to consider for their potential externalities. The first is the fact that increasing ρ causes previously-indifferent organizations ($\bar{y} = \rho + \delta$) to bunch. This affects these organizations' probability of bunching again in the future, though there is no effect on going missing in the future because by definition these are organizations with $\mu > 0$, else they would go missing when $\bar{y} = \rho + \delta$. The second effect is that some organizations that would go missing will instead bunch, making them candidates to instead go missing the following year. For simplicity I will ignore the second effect because it will involve very few organizations and will result simply in retiming of when the organization goes missing. With this simplification, the information cost of increasing ρ is

$$\frac{d}{d\rho}\prod(\rho) = 1\{\mu \geq 0\}\,\pi\sum_{t=0}^{\infty}\beta^t\left[-1-\sum_{s=1}^{\infty}\beta^s\eta\right] f(\rho+\delta|\bar{y}_{t-1}, y_{t-1})$$

$$= -\pi\cdot 1\{\mu \geq 0\}\, f(\rho+\delta|\bar{y}_{t-1}, y_{t-1})\,\frac{1-\beta+\eta\beta}{(1-\beta)^2}$$

Combining the direct effect on charities and the indirect effect on the number of forms gives the total effect on social welfare. Assuming that the income distribution is not changing over time and that the cost of missing η is uncorrelated with $\bar{y}$ gives

$$\frac{dW}{d\rho} = \int\left[\frac{d}{d\rho}V(\rho,\phi,\gamma) + \frac{d}{d\rho}\prod(\rho)\right] g(\gamma)\,d\gamma$$

$$= (1-H(\triangle))\,\frac{1}{1-\beta}\,(F(\rho+\delta)-F(\rho))\,E\left[C_a(\hat{y},\hat{y}-\rho,\gamma,\bar{y})\,|\rho<\bar{y}\leq\rho+\delta\right] - \pi\,(1-H(0))\,f(\rho+\delta)\,\frac{1-\beta+\eta\beta}{(1-\beta)^2}$$

$$\frac{dW}{d\rho}\geq 0 \Leftrightarrow \frac{\pi}{E\left[C_a(\hat{y},\hat{y}-\rho,\gamma,\bar{y})\,|\rho<\bar{y}\leq\rho+\delta\right]} \leq \frac{1-\beta}{1-\beta+\eta\beta}\cdot\frac{1-H(\triangle)}{1-H(0)}\cdot\frac{\left(F\left(\rho+\bar{\delta}\right)-F(\rho)\right)}{f\left(\rho+\bar{\delta}\right)} \tag{2}$$

Comparison of 2 with 1 (with the simplifying assumption $b = 1$) shows how the dynamics and missing observations affect the welfare analysis. The left-hand side of the inequality is still the ratio of social benefits from forms filed to the marginal cost of bunching, and the right-hand side still includes the bunching ratio. The two new terms on the right-hand side highlight the importance of dynamics and missing observations, respectively. First, the right-hand side is decreasing in η, the persistence of bunching. If bunching is highly persistent then there will be a more narrow range of values of π for which the government should raise the level of the notch because the marginal bunching this will induce will have persistent effects and hence a greater cost in present value. The right-hand side is increasing in $\frac{1-H(\triangle)}{1-H(0)}$, the ratio of non-missing organizations in the reduced range vs. at higher income levels where bunching is not a desirable option. If organizations are more likely to go missing when above the reduced range than when they are when in it then there will be a wider range of values of π for which the government should raise the level of the notch because it would benefit many non-missing charities in the reduced range and induce fewer non-missing charities just above the reduced range to bunch.

4 Empirical Analysis

4.1 Static Estimates and Non-causal Evidence of Repeated Bunching

For completeness, I first present results of standard techniques. Figure 3 (also presented by Marx (2018)) provides an example of the standard approach to bunching estimation. Income is binned, and the counterfactual income distribution is estimated as a polynomial through bin counts outside of the omitted range around the notch. Compared to this counterfactual, the actual bin counts reveal excess mass just below the notch and a reduction just above. This shift can be interpreted as a response by charities because unilaterally determine how much income to report, unlike labor supply outcomes that may depend on strategic interactions of workers and firms (Chetty et al., 2011).

Estimation results will be provided later for comparison with the dynamic bunching estimates. Marx (2018) shows that the estimates from the standard approach are not reliable. One reason for this is extensive-margin responses, which are assumed away by the standard approach. Consistent with extensive-margin responses, the estimated reduced mass above the notch is significantly larger than the estimated excess below the notch, consistent with the extensive-margin responses estimated later. Another concern is that charities faced the same notch every year, which Marx (2018) shows to be problematic for bunching estimation if bunching in one year reduces income in later years. One way that this can happen is if agents bunch repeatedly.

A common technique with panel data can provide suggestive evidence of repeat bunching. If charities bunch repeatedly then those currently in the bunching range should have a heightened probability of remaining in that range. Figure 4 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 an interpolated counterfactual prediction of only 5 percent.

To estimate repeated bunching within the static framework I again construct bins of current receipts. I then estimate the probability that in h years an organization remains in its current bin as

$$p_{i,t+h} = \beta \cdot bunchbin_i + \sum_{k=1}^{K} \alpha_k r_{it}^k + \gamma_t + e_{it} \tag{3}$$

where $p_{i,t+h} = 1\left[r_{i,t+h} = r_{i,t}\right]$ is an indicator for remaining in the same bin h years in the future, the width of each bin is bw and $bunchbin_i = 1\{r_{it} \in [notch - bw, notch)\}$ is an indicator for having current receipts in the bunching range, $\sum_{k=1}^{K} \alpha_k r_{it}^k$ is again a polynomial in receipts that provides the counterfactual for the bunching range, and γ_t is a vector of year dummies.[13]

Table 3 reports the results from estimating regression 3 for horizons up to 10 years. I find that charities bunch just below the reporting notch for many 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.

Evidence that charities in the bunching region have a heightened probability of remaining there does not indicate whether they remain because they continue to bunch or because it is slow-growing charities that bunch in the first place. Similarly, one might hypothesize that low-expense charities are most likely to bunch and that when charities bunch they must reduce their expenses to balance the budget. Alternative methods are needed to estimate the causal effect of approaching the notch and to separately estimate heterogeneity in the bunching response and concurrent responses in other variables that may correlate with bunching.

4.2 Temporary Notch – Permanent Effects on Growth

Variation in the IRS reporting notch for public charities makes it possible to estimate effects of a temporary notch. After decades with a notch at a nominal value of \$100,000, the notch was moved to \$1,000,000 for 2008, \$500,000 for 2009, and \$200,000 thereafter. The 2008 and 2009 notches were therefore only temporary (one-year) phenomena that should not have induced repeated bunching or manipulation at incomes far from the level of the notch. I exploit this temporary nature, focusing on the 2009 notch, which fell at an income level with many more charities.

Denote by $near2009notch_i$ an indicator for the treatment group of charities in the bin of 2009 log receipts that straddles the notch. I assign each charity i to an equally-sized bin of 2009 log receipts $bin2009_i$ and estimate simple equations of the form

$$E\left(Y_{it} | bin2009_i\right) = \beta \cdot near2009notch_i + \sum_{j=1}^{J} \alpha_j bin2009_i^j \tag{4}$$

The second term is a polynomial in bin level that provides a counterfactual for the level that should be

[13]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.

expected in the $near2009notch_i$ bin.[14] The parameter β therefore describes the deviation of the average deviation from conditional expectation among charities in the omitted region around the notch. Outcomes Y_{it} can include functions of receipts in 2009 to identify and measure the bunching response, but also variables measured in earlier years as a test of the assumption that charities *near* the 2009 notch are not a selected sample, as well as variables measured in later years to test for long-term effects of one-time bunching.

I examine three different measures of income. The first is simply the level of (log) receipts, which is expected to be negatively affected in 2009 by the opportunity to bunch. The second, labeled "Cross 2009 Threshold," is an indicator for whether the organization's current receipts are above the level corresponding to the notch. This dummy variable indicates that the charity has receipts greater than .065 plus the level of the minimum income in the bin that it occupies in 2009, since the minimum income in the treatment bin is .065 log points below the notch. Bunching should also have a negative effect on the probability that treated observations cross the 2009 threshold in 2009. Finally, I construct an indicator for current receipts that lie within the same bin that the charity occupies in 2009. This dummy takes the value of one for all observations in 2009 but provides another useful measure of pre-trends or long-term effects.

Figure 5 plots the results of one of the regressions. Data are binned by log receipts in 2009 (relative to the temporary notch), and the outcome is the share of organizations in each bin that "Cross 2009 Threshold" in 2010. The counterfactual estimated from the control bins away from the notch indicates that close to 54 percent of the organizations in the treatment bin straddling the temporary notch should have had receipts above $500,000 in 2010. The share that actually achieved this size was closer to 52 percent, and confidence intervals on the counterfactual and observed share indicate that this result is statistically significant.

Results of the temporary-notch regressions appear in Table 4 (also presented by Marx (2018)). Each cell of the table reports the estimate of the parameter β for a different outcome and year. In the first row the outcome is log receipts, which is negatively impacted in 2009, as expected. The coefficient for this year implies that the average charity near the notch reduces income by .003 log points, or roughly $1500 ($500,000*.003). Estimated deviations from expectation are negative for subsequent years, but standard errors are large because, as will be seen later, the distribution of year-over-year income growth has fat tails. The alternatives income measures, which focus on more central growth rates, can therefore offer more-precisely-estimated evidence. In the second row, one can see that the probability of being above the notch in 2009 is reduced by 10.1 percentage points among the treatment group. These charities are not significantly

[14] The results depicted in this section are obtained with a simple quadratic function of bin level, and range starting at receipts greater than $100,000 to avoid lingering effects of the original notch, and log receipts bins of width .155, the width of the smallest omitted range that provides robust static bunching estimates. This approach is still susceptible to the problem of missing mass due to systematic attrition of organizations crossing the notch. However, for the temporary notch of 2009 this does appear to be a problem: the excess mass below the notch estimated using the static method is not statistically different from the reduction in mass above the notch or from the dynamic estimate obtained in Section 4.4, and the dynamic estimate of extensive-margin responses is not significantly different from zero.

different in this regard prior to 2009, but bunching in 2009 appears to cause a permanent 2 percentage-point reduction in the probability that treated charities ever achieve receipts greater than $500,000. Similarly, the third row shows that while these charities were no more or less likely to be in their 2009 income bin in years before 2009, the probability that they remain in this bin (rather than growing out of it) is permanently increased by at least .4 percentage points.

Permanent effects of a temporary notch provide strong evidence that manipulation is not entirely carried out by misreporting income. Simple underreporting of income would have no effect on income in later years, when there was no reason to underreport, and retiming of income would actually increase earnings in 2010. The results are not sufficiently precise to fully rule out some misreporting, but there does appear to be some permanent, monetary cost associated with bunching. The finding of the temporary notch affecting real income is consistent with the evidence that follows for the permanent notch that was in place for all sample years before 2008.

4.3 Dynamic Ordinary Least Squares – Repeated Bunching and Heterogeneity

The OLS dynamic bunching design shows that 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 has long-run effects on growth.

Figure 6 provides visual evidence supporting the dynamic identification strategy. As was the case in the PSID data, growth distributions do not vary dramatically across levels of base-year receipts. Since the notch falls at different levels of growth for these different levels of base-year receipts, the distortions induced by the notch can be identified by comparison with growth distributions from other base-year levels.

Regression analysis shows highly significant manipulation of receipts.[15] In column (1) of Table 5 we see that receipt growth of charities nearing the notch is reduced by .0017 log points. The average reduction is therefore about 0.17 percent 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 four 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

[15] I present specifications with a full set of growth range dummies and the interactions of each growth rate bin with a quadratic function of income (so that $J = 2$ and $K = 2$ in the estimating equation), though results are robust to alternative choices. Details of my implementation of the dynamic OLS estimates appear in Online Appendix A.

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 (also presented by Marx (2018)) 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 question 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."[16] 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.

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 heterogeneity by size presented in Table 8). 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

[16] 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.

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.

4.4 Dynamic Maximum Likelihood Estimation – Bunching and Extensive-Margin Responses

I estimate several quantities using maximum likelihood. The bunching share b 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 also 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 the notch. Finally, I report the extent of the extensive-margin response, which I calculate separately for charities that should grow to the range of income from which some choose to bunch and for incomes above this range.

Attrition is common in the charity data. Attrition could be due to late filing, earning income below the level at which filing is required, shutting down, merging, or simply non-compliance. I estimate three types of attrition. First, I include polynomials in base-year income below and above the notch 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 agents with income below $25,000 do not have to report. Third, I allow for the extensive-margin responses that reduce the share of observations that grow to income levels above the notch. Further details of the estimation are provided in Online Appendix B.

Figure 9 illustrates the approach as well as the goodness of fit. The figure is populated using an illustrative group of charities with base-year receipts just below the notch. The empirical distribution of growth rates is non-Normal, with a sharper peak around zero growth and fatter tales, thus requiring a flexible approach to fit. Fitting is done with observations outside of the omitted range of growth rates that would bring these charities close to the notch. The fit is good outside of the omitted range, while inside the range the bunching response leads to a shift of mass from positive growth rates that would cause charities to cross the notch to nonpositive rates that keep them below it. The extent of this downward shift identifies the bunching response.

Table 9 displays maximum likelihood dynamic bunching estimates. The first column displays the static bunching estimates for comparison. The second and third columns, respectively, contain dynamic bunching estimates for the years when the reporting threshold was $100,000 and the years when the reporting threshold was at the current level of $200,000. Extensive-margin responses are significant and help explain the difference between the static estimates of the excess mass below the notch and the reduced mass above the notch.

The top panel of Table 9 shows dynamic bunching estimates of parameters governing bunching and systematic attrition. The first parameter estimate gives the bunching propensity among charities that have current, base-year receipts below the notch. For the current notch, an estimated 4.8 percent of the charities that should grow from an income level below the notch to a new level above it in each year 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 still significant but considerably smaller for both the past and current notch. 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 infrequency of bunching by charities coming from above the notch is consistent with (unreported) results from OLS dynamic estimation.

The third and fourth rows in the table show that attrition is significantly related to the notch. The estimated parameters indicate the share of charities that should be crossing the notch from below but instead go missing from the sample. To allow for flexibility and avoid bias in the bunching estimates, the attrition rate is estimated separately for those that should barely cross the notch (into the reduced range) and those that should grow to an even higher level of income. Both estimates are economically and statistically significant for the historical notch, with close to ten percent of these growing charities leaving the sample. Comparing the attrition and bunching propensities, a combined 10.6 percent of charities avoid filing when first crossing the notch to the reduced range, and the number of charities doing so by bunching is dwarfed by the number responding on the extensive margin. Bunching became more common and extensive-margin responses became less common when the reporting threshold increased to $200,000, perhaps because attaining higher incomes levels increases the perceived risk of not reporting. The static approach, which assumes away extensive-margin responses, does not provide estimates of any of these parameters in the top panel of the table.

The lower panel of Table 9 reveals the estimated excess share of charities below the notch and reduction in the share above it. The excess and reduction are found by aggregating the bunching and attrition propensities across all observations in the base year according to their counterfactual probability of moving to the reduced region in the next year. According to the dynamic bunching estimates for the $100,000 notch, the share of charities that appeared in the bunching range exceeded by .103 percentage points the quantity that would have been found in the absence of a response. Consistent with the difference in estimated magnitudes of the propensities to bunch or leave the sample, the reduction of mass in the range just above the notch is significantly greater at .366 percent. Following the static approach with observations in the same years gives estimates of the excess and reduced mass that are much closer to each other and both significantly different from the dynamic estimates. Intuitively, the extensive-margin responses and persistence of both income and bunching lead to a reduction in mass everywhere above the notch, not just in the reduced range. The effect on the density above the reduced range introduces bias into the static estimate of the counterfactual, which assumes that the density above the reduced range is unaffected. The excess mass just below the $200,000 notch is comparable to that at the old notch. The decline in extensive-margin responses has brought the reduction of mass from the range just above the notch closer to the level of excess mass below it, but the reduction remains significantly greater.

Finally, the bunching ratio gives the ratio of excess bunching mass to the counterfactual density at the notch, here reported for the density in levels so that the ratio can be interpreted as a dollar amount. Standard, static bunching estimation would lead the researcher to conlcude that the average charity reduces income by $753 to bunch to avoid filing Form 990. In contrast, the dynamic estimate for the $100,000 notch indicates that the average charity was only willing to reduce income by about $400. Compared to this dynamic estimate, the static estimates of the excess mass and bunching ratio are biased upwards by roughly ninety percent. The final column of the table shows the dynamic estimate of the amount of bunching when the notch moved to $200,000. The excess mass remained about the same despite a much smaller density at the new level of the notch, and so the estimated bunching ratio doubled to $934. The difference may be due in part to differences between charities at different income levels, but one would also expect that charities should now be willing to reduce income by an even greater amount, given that the notch was moved to a higher income level precisely because the form was made more onerous.

4.5 Donut RD – Duration of Extensive-Margin Responses

For welfare calculations it is also important to know whether charities exit permanently or simply delay filing until they are able to comply. Table 10 provides reduced-form evidence on this question. In these regressions, attrition variables are regressed on flexible functions of gross receipts to examine the duration of missing-data spells among charities that are just below the notch and hence likely to cross it and go missing. I employ the donut regression discontinuity approach used by Barreca et al. (2011), excluding observations in the bunching range and in an equally-sized range above the notch because it is mostly likely that bunching is correlating with the probability of leaving the sample. For each outcome variable, results are provided for regressions with third-, fourth-, and fifth-order polynomials on either side of the notch, and the specification that minimizes the Akaike Information Criterion is listed in bold. Results for the first outcome are consistent with the maximum likelihood estimation results: The probability of going missing from the sample in the next year is elevated by about 2.3 percentage points for charities with receipts just below the notch. The difference in the probability of permanent exit, however, is estimated to be less than one tenth of one percent and statistically insignificant. It appears, therefore, that charities crossing the notch go missing temporarily but eventually return to the sample. Consistent with this interpretation, the estimated number of years missing from the sample is elevated by .035. Dividing this estimate by the 2.3 percentage-point difference in the share going missing indicates that extensive-margin responders are leaving the sample for 1.5 years on average.

4.6 Welfare Implications of Parameter Estimates

The dynamic bunching methodology provides estimates of the key parameters for evaluating the welfare consequences of adjusting the notch. These key parameters, as shown in 2, include the increased likelihood of bunching in future years if bunching now (η), the share of observations missing from the reduced region ($H(\triangle)$), the share of observations missing from above the reduced region ($H(0)$), and the bunching ratio of excess mass to density at the notch. The bunching ratios for the historical notch and the current notch have already been shown in Table 9, as have the shares of missing observations. Because the attrition rate for charities moving to the reduced region was not very different from the attrition rate among those moving to higher incomes, the adjustment for missing observations turns out not to affect the welfare calculation much. In contrast, recognizing the persistence of bunching makes a considerable difference. For my estimate of η I use the effect of the temporary \$500,000 notch on the probability of crossing \$500,000 in the future, which was shown in Table 4. I take the reduced fraction of charities crossing the threshold in the future (using the midpoint of .019 between the three estimates that vary from .018 to .020) and divide it by the fraction induced not to cross in the year of the notch (.101) to obtain the effect per buncher, and I arrive an estimate of $\hat{\eta} = .188$. Plugging the historical-notch estimates from Table 9 into 2 and assuming $\beta = .97$ gives

$$\frac{dW}{d\rho} \geq 0 \Leftrightarrow \frac{\pi}{E\left[C_a\left(\hat{y}, \hat{y}-\rho, \gamma, \bar{y}\right) | \rho < \bar{y} \leq \rho + \delta\right]} \leq 54.34$$

Because $E\left[C_a\left(\hat{y}, \hat{y}-\rho, \gamma, \bar{y}\right) | \rho < \bar{y} \leq \rho + \delta\right] \in (0,1)$, the left-hand side has value at least equal to π, the social value (net of government administrative cost) of a completed long form relative to an EZ form. Thus, $\pi \geq 54.34$ would have been a sufficient condition to *lower* the income level at which more reporting was required. A static analysis would have concluded that the appropriate threshold would be at the static bunching ratio of 753.21, a required social benefit nearly fifteen times greater than that estimated with the dynamic methodology.

For the current notch I find

$$\frac{dW}{d\rho} \geq 0 \Leftrightarrow \frac{\pi}{E\left[C_a\left(\hat{y}, \hat{y}-\rho, \gamma, \bar{y}\right)|\rho < \bar{y} \leq \rho+\delta\right]} \leq 120.71$$

The threshold value required to lower the notch has risen because bunching has increased with the introduction of a more complicated form. Because empirical testing rejected the possibility that bunching was simply misreporting, $E\left[C_a\left(\hat{y}, \hat{y}-\rho, \gamma, \bar{y}\right)|\rho < \bar{y} \leq \rho+\delta\right]$ appears to be significantly greater than 0 and perhaps close to 1. If we assume the marginal cost of avoidance is 1, then increasing the income threshold, say by indexing it to inflation, would be beneficial if and only if the marginal social value of a long form is less than or equal to \$120.71. I leave estimation of the social value of Form 990, which would require information about the auditing processes of the Internal Revenue Service, to future work.

5 Conclusion

This paper estimates the compliance cost of a financial reporting requirement for U.S. charities. The average charity is currently willing to reduce income by about \$900 to avoid reporting more information, and many leave the sample in the first year that their reporting burden increases. The fact that small charities who had not previously filed the simplified form were most likely to manipulate income provides evidence that much of the compliance cost is a one-time adjustment, and bunching to avoid this cost appears to be driven not simply by misreporting but rather by foregoing income in a way that permanently reduces growth.

The results of this paper provide new evidence on compliance costs and the growth of firms. Firms have been found to bunch at regulatory thresholds in some settings (e.g. Garicano et al., 2016; Gourio and Roys, 2014) and not others (Hsieh and Olken, 2014), perhaps due to heterogeneity that the dynamic bunching methodology can help to describe. For example, compliance costs appear to have an important fixed component because their burden is proportionately heavier on smaller businesses (Slemrod and Venkatesh, 2002). Public charities also appear to face scale economies in compliance, which consumes 7 percent of the annual budgets of surveyed charities with revenue below \$100,000 (Blumenthal and Kalambokidis, 2006b). The heterogeneity results in this paper reinforce this evidence of economies of scale in compliance costs and indicate that such economies may obtain in terms of staffing or assets even when income is held constant.

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.

Barreca, Alan I., Melanie Guldi, Jason M. Lindo, and Glen R. Waddell, "Saving Babies? Revisiting the effect of very low birth weight classification," *The Quarterly Journal of Economics*, 2011, *123* (4), 2117–2123.

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

_ **and** _ , "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.

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.

Fack, Gabrielle and Camille Landais, “The effect of Tax Enforcement on Tax elasticities: Evidence from Charitable Contributions in France,” *Journal of Public Economics*, 2016, *133*, 23–40.

_ **and** _ , *Philanthropy, Tax Policy and Tax Cheating: A Long Run Perspective on US Data*, Oxford University Press, 2016.

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.

Garicano, Luis, Claire Lelarge, and John Van Reenen, “Firm Size Distortions and the Productivity Distribution: Evidence from France,” *American Economic Review*, 2016, (106), 3439–3479.

Gourio, François and Nicolas A. Roys, “Size-Dependent Regulations, Firm Size Distribution, and Reallocation,” *Quantitative Economics*, 2014, (5), 377–416.

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.

Hsieh, Chang-Tai and Benjamin A. Olken, “The Missing ”Missing Middle”,” *Journal of Economic Perspectives*, 2014, *28* (3), 89–108.

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 Public Economics*, 2004, *88*, 559–576.

Kleven, Henrik J. and Mazhar Waseem, “Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan,” *Quarterly Journal of Economics*, September 2013, *128(2)*, 669–723.

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

Kline, Patrick and Melissa Tartari, “Bounding the Labor Supply Responses to a Randomized Welfare Experiment: A Revealed Preference Approach,” *NBER Working Paper No. 20838*, January 2015.

Marx, Benjamin M., "Has Regulation of Charitable Foundations Thrown the Baby Out With the Bath Water?," *Journal of Public Economics*, September 2015, *129*, 63–76.

— , "Dynamic Bunching Estimation with Panel Data," *MPRE Paper 88647*, 2018.

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.

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

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.

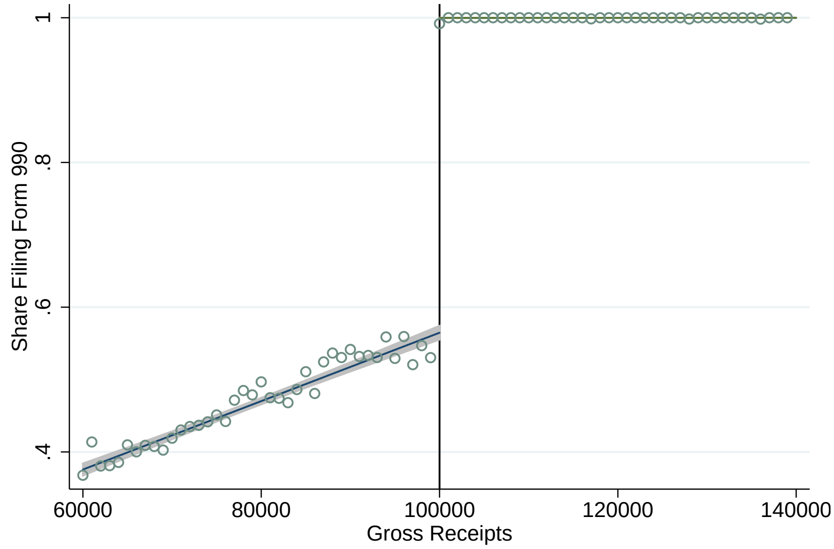
— **and Varsha Venkatesh**, "The Income Tax Compliance Costs of Large and Mid-Sized Businesses," Technical Report, Report to the Internal Revenue SErvice Large and Mid-Size Business Division 2002.

St.Clair, Travis, "How Do Nonprofits Respond to Regulatory Thresholds: Evidence From New York's Audit Requirements," *Journal of Policy Analysis and Management*, 2016, *35*, 772–790.

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.

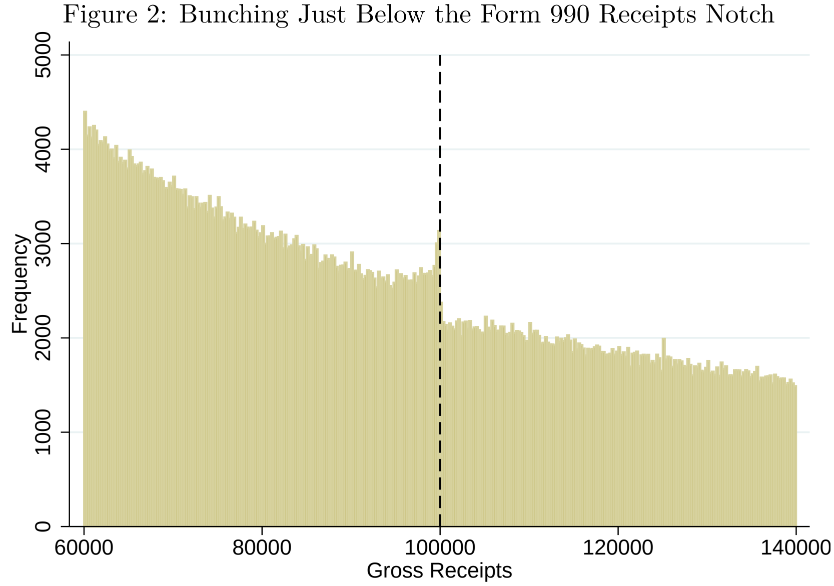
Figures and Tables

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



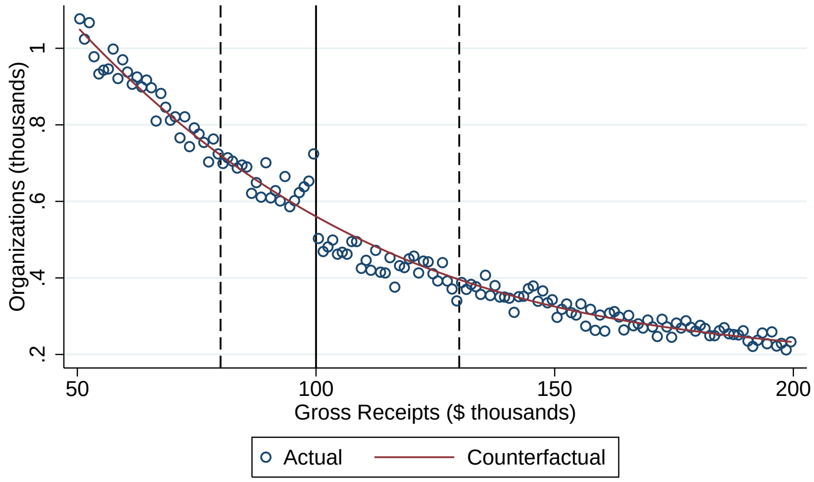
Notes: The figure shows the probability of filing Form 990 (vs. 990-EZ) in 2007 as a function of gross receipts below and above the $100,000 notch at which charities lose eligibility to instead file Form 990-EZ. Circles show the mean within each $1000 receipts bin, and curves with standard error bands show the results of a linear regression for each side of the notch. The share of organizations filing the longer form is increasing in receipts up to the notch, then jumps as expected, with nearly 100% compliance above the notch. N=72,354.

Figure 2: Bunching Just Below the Form 990 Receipts Notch



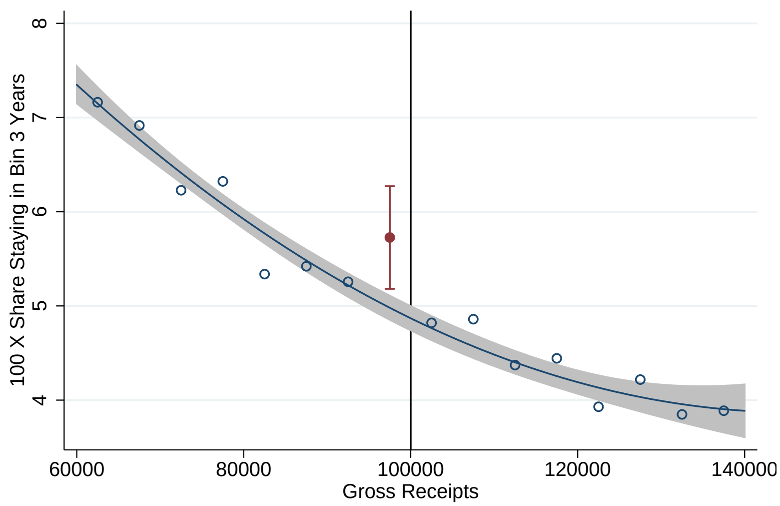
Notes: The figure is a histogram of gross receipts. An excess of charities just below the $100,000 notch appears as bunching in what is otherwise a smooth distribution. N=810,869. Bin width=$250. Years 1999-2007 pooled.

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



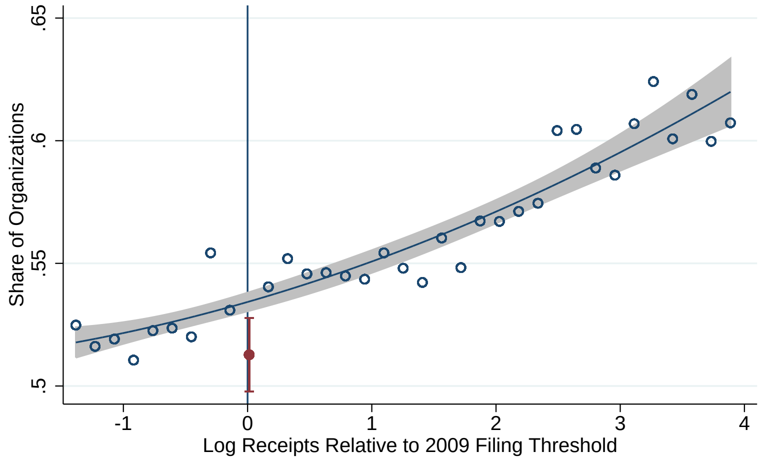
Notes: The figure shows the deviation of the 2006 distribution, represented by a histogram in blue circles, from a smooth counterfactual. Each bin is treated as an observation. Bin counts are regressed on a polynomial of degree 3, which estimates the counterfactual distribution, and a dummy variable 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 $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: Repeated Bunching: Share of Charities Staying Within Bin For 3 Years



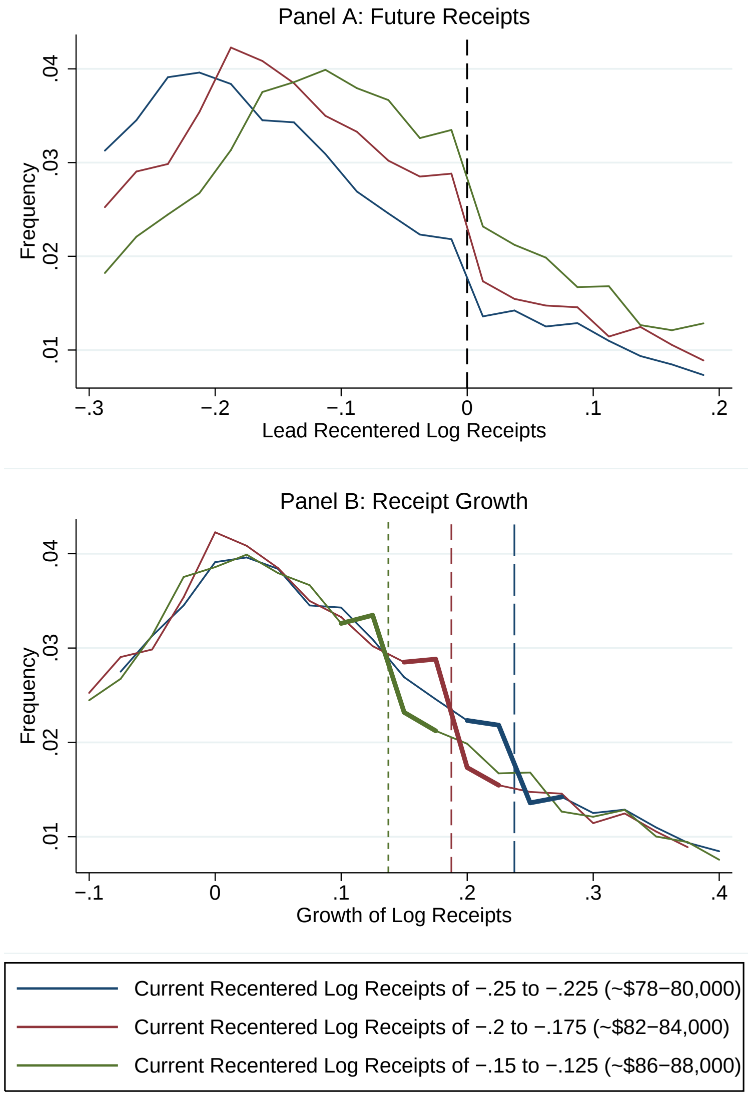
Notes: The figure shows the share of charities in each $5000 bin of current gross receipts that remains in the same bin 3 years later. The curve represents a quadratic fit to these probabilities for bins other than the bin just below the reporting notch. The marker with a 95-percent confidence interval shows that organizations in the bunching region just below the notch are especially likely to remain at their current income level for several years. Standard errors clustered by state. N=329,448. Bin width=$5000.

Figure 5: Lasting Effects of a Temporary Notch on Future Income of Charities



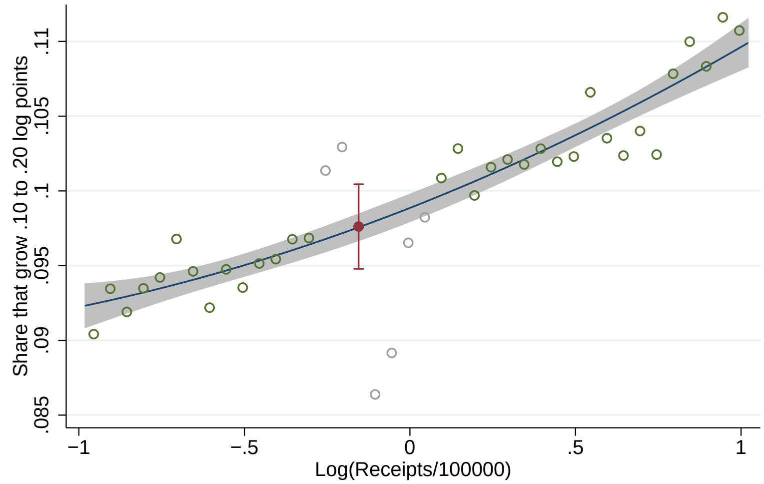
Notes: The figure shows how a measure of 2010 income ("Cross 2009 Threshold" as described in the text) varies with 2009 income. The marker with a 95-percent confidence interval shows that of organizations with 2009 receipts close to that year's temporary notch of $500,000 roughly 52 percent had incomes over $500,000 in 2010. This percentage is significantly less than the counterfactual interpolated from the corresponding shares for charities in other 2009-receipts bins. Because the treatment bin straddles the temporary 2009 notch and includes both charities that bunched in 2009 and those that did not, the difference between the solid marker and the counterfactual provides an estimate of the causal effect of having income near the temporary notch. N=127,855. Treatment range is -.065 to .09. Bin width = .155.

Figure 6: Distorted and Undistorted Sections of Conditional Distributions of Charities' Future Receipts/Growth



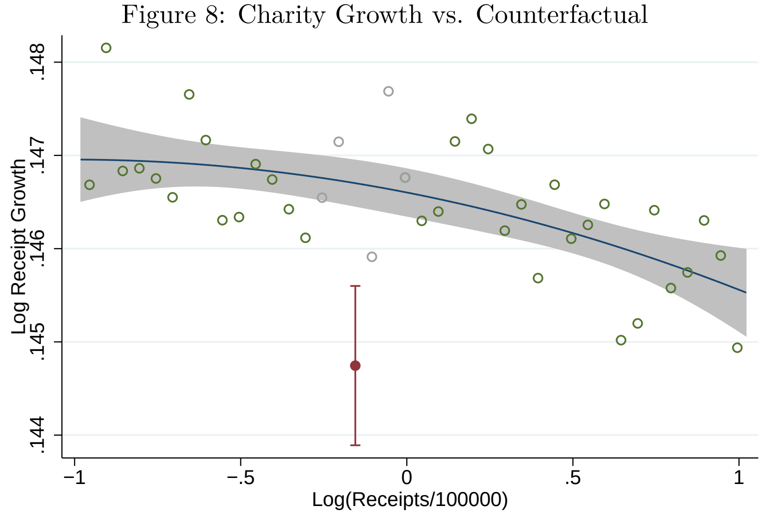
Notes: The figure shows the distribution of future receipts (Panel A) and growth to future receipts (Panel B) for charities in three illustrative bins of current receipts. The distributions for each group exhibit a spike at incomes just below the notch and a depression just above it, indicating manipulation 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 7: Share Growing To a Range That Spans the Notch is Unaffected



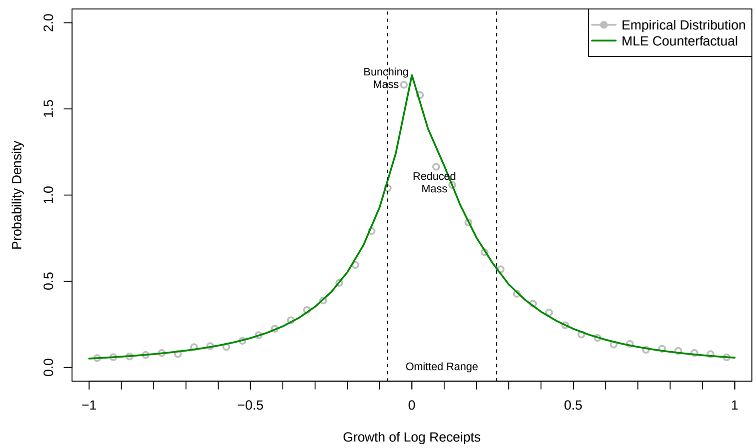
Notes: The figure shows the results of regressing the probability of growing log receipts by .1 to .2 (from the current year to the next) on a quadratic in current recentered log receipts and a dummy ("Near Notch") for the bin for which future receipts lie in the "omitted range" straddling the notch. 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 8: Charity Growth vs. Counterfactual



Notes: The figure shows growth of income from the current year to the next year as a function of current income (recentered around the reporting notch at $100,000). The figure sample consists of organizations in an illustrative growth bin that includes are organizations with growth between .1 and .2 log points. The marker with a 95-percent confidence interval represents the bin (defining the "$nearnotch_{it}$" dummy described in the text) for which growth of .1 to .2 implies that future receipts lie in the "omitted range" straddling the notch. The conditional average growth rate of these charities is just below .145, which is significantly less than the counterfactual growth rate interpolated from charities with higher and lower current incomes. The difference is interpreted as a measure bunching; some charities that approach the notch reduce their income to stay below it, and therefore conditional average growth is less than predicted. N=152,191. Omitted range is -.08 to .07. Bin width = .05.

Figure 9: Estimation of the Distribution of Charities' Growth Rates



Notes: The figure shows the observed and estimated counterfactual densities of growth in log receipts conditional on current income for an illustrative group of charities. The curve shows the fit of the maximum likelihood estimate of this distribution to the data represented by the histogram of circular markers. The sample consists of organizations starting from a range of income levels .0 to .025 log points below the notch, those that would cross the notch in the next year if growing at a rate in excess of .025. A range of growth rates around 0-.025 is omitted from estimation of the counterfactual. The counterfactual therefore fits the observed distribution closely except that bunching increases the mass at growth rates near the bottom of the omitted range and increases the mass at higher incomes within the omitted range. Bin width = .05. N=2,815,026.

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: 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. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Persistent Reductions in Charities' Incomes After the One-Time Notch of 2009

	2007	2008	2009	2010	2011	2012
Log Receipts	-0.001	0.009	-0.003	-0.007	-0.013	-0.018
	(0.010)	(0.008)	(0.001)***	(0.008)	(0.009)	(0.012)
Cross 2009 Threshold	0.003	0.006	-0.104	-0.022	-0.020	-0.018
	(0.008)	(0.008)	(0.007)***	(0.008)***	(0.008)**	(0.010)*
Same Receipts as in 2009	-0.000	0.001		0.005	0.004	0.007
	(0.001)	(0.002)		(0.002)**	(0.002)*	(0.003)***
Observations	48,716	107,579	105,160	127,855	115,601	104,760

Notes: The table shows the results of regressions of three different variables on a quadratic function of binned log receipts in 2009 and a "Near Notch" dummy for the bin that straddles the $500,000 notch for that year. The table shows the estimate of the coefficient on the "Near Notch" dummy, which represents the causal effect of having income near the notch in 2009. The first row shows that log receipts of charities near the threshold in 2009 are significantly lower than expected in that year, as expected due to bunching. Point estimates remain negative in subsequent years but standard errors are large. The outcome in the second row is a dummy for crossing the level of growth corresponding to the notch ("Cross" as defined in the text). The coefficients indicate charities experience a significant, permanent reduction of at least one percentage point in the probability of having income over $500,000 in any year after 2009. The outcome for the third row is an indicator equal to one if the charity is in the same log receipts bin that it is in 2009, and the results indicate that charities are significantly less likely to have grown out of their bin in 2009. In years before 2009 there are no significant differences between the treated charities and the interpolated counterfactual. Robust Huber-White standard errors are displayed. Bins have width .155 and extend from 1.615 log points below the notch to 3.19 log points above it (roughly $100,000 to $12 mil). These parameters give 35 control bins in addition to the treatment bin, and bunching estimates are robust to changes in these parameter choices. *** p<0.01, ** p<0.05, * p<0.1

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)
Cross			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 the notch (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, are not precisely estimated. The sample includes all charities growing by 0 to 1 log points. Standard errors are clustered by state. *** p<0.01, ** p<0.05, * p<0.1

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 significant 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. *** p<0.01, ** p<0.05, * p<0.1

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

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

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

	Static 1991-2007	Dynamic 1991-2007	Dynamic 2010 & 2011
Share bunching from below notch		0.026***	0.048***
		(.003)	(.012)
Share bunching from above notch		0.005***	0.007**
		(.001)	(.003)
Attrition of those crossing to reduced range		0.080***	0.036***
		(.004)	(.011)
Attrition of those crossing to higher incomes		0.093***	0.001
		(.005)	(.011)
Excess mass just below the notch (*100)	.194***	0.103***	.105***
	(.017)	(.007)	(.017)
Reduction in mass in reduced range (*100)	.293***	0.354***	.193***
	(.019)	(.012)	(.025)
Bunching ratio	753.21***	404.92***	933.99***
	(67.04)	(65.14)	(150.56)
N	2,196,564	2,815,026	386,805

Notes: The table shows the results of maximum likelihood dynamic bunching estimation, along with static estimates on similar sample for comparison. The top panel of the figure provides estimates from the dynamic design that cannot be obtained from the static approach. 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. The next two parameter estimates imply that significant share of the charities with current income below the notch should grow to an income level above the notch but instead exit from the sample. The lower panel shows that the static approach overestimates the excess number of organizations just below the notch and underestimates the number that should be just above it. All regressions allow for attrition that can vary with current income as described in the text. The sample size for static estimation is smaller than that for dynamic because the latter includes all charities appearing in the base year while the former excludes charities that were missing or far above the notch in the next year, but static estimates are rescaled to have the same denominator as the dynamic estimates for comparability. Standard errors for dynamic estimates are calculated numerically. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Donut RD Estimates of Propensities to Leave the Sample Temporarily vs. Permanently

	Goes Missing	Goes Missing	Goes Missing	Exits	Exits	Exits	Years Missing	Years Missing	Years Missing
Receipts < $100,000 notch	0.009	**0.023**	0.027	**0.000**	0.002	0.007	0.015	**0.035**	0.027
	(0.003)***	**(0.005)*** **	(0.008)***	**(0.002)**	(0.003)	(0.005)	(0.007)**	**(0.010)*** **	(0.016)*
Degree of Polynomials	3	**4**	5	**3**	4	5	3	**4**	5
Akaike Information Criterion	1,772,144	**1,772,138**	1,772,141	**-226,848**	-226,844	-226,845	3,881,354	**3,881,352**	3,881,355

Notes: The table shows the results of donut regression discontinuity estimation of the probability of leaving the sample conditional on current gross receipts. Regressions include observations within one log point of the $100,000 receipts notch but not within .05 log points, where bunching and hence sample selection are greatest. Regressions include a polynomial for observations on each side of the notch, and for each outcome variable the specification with the polynomial order that minimizes the Akaike Information Criterion is shown in bold. Results for the three outcomes indicate, respectively, that charities with incomes below the notch are significantly more likely to go missing from the sample in the next year, not significantly more likely to exit permanently, but rather to have significantly more years pass before the charity is observed again. Robust standard errors are reported in parentheses. N=1,765,089. *** p<0.01, ** p<0.05, * p<0.1

Appendix A - Details of Dynamic Ordinary Least Squares Estimation

Section 4.3 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.

The estimating equation for outcome Y_{it+1} is

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

where r_{it} is current recentered log gross receipts and $D\left(a\right)_{it+1} = 1\left\{r_{it+1} - r_{it} \in [a, a+w)\right\}$ is an indicator for growth falling within one of the A growth rate ranges of width w. 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 ($K = 0$). I present specifications with $J = 2$ and $K = 2$; results are highly robust to different specifications.

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 4.4, 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.

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.

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 B - Details of Dynamic Maximum Likelihood Estimation

Section 4.4 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

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

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

$$+ \left(\pi_3^l + \tau_3^l r\right)\left[\exp\left(-(g-\theta)^2\right) - 1\right] + \left(\pi_4^l + \tau_4^l r\right)\arctan\left(\left(\pi_5^l + \tau_5^l r\right)(g-\theta)\right)$$

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

$$+ \left(\pi_3^u + \tau_3^u r\right)\left[\exp\left(-(g-\theta)^2\right) - 1\right] + \left(\pi_4^u + \tau_4^u r\right)\arctan\left(\left(\pi_5^u + \tau_5^u r\right)(g-\theta)\right)$$

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

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

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

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

The implied latent density is

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

where $P_l'(g, r, \theta) =$ and $P_u'(g, r, \theta)$ 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 $notch := \log(100,000)$ 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.[17] Since the empirical distribution has fat tails, with observed growth rates of absolute

[17] Missing and bunching observations could be assigned to any value of g^*. Identification uses the count of missing and the

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 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 $rmin - r$.[18] Second, some share $\lambda(r)$ 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 $\delta(r, g)$ 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 $(rmin - r)$. When estimating these attrition functions I allow $\lambda(r)$ to be quadratic in r below the notch and linear above it, with different intercepts for each. The share exiting instead of crossing the notch, $\delta(r, g)$, is only relevant for those with r below the notch and is allowed to take different values depending on whether growth takes the observation to the reduced range, where bunching offers an alternative to going missing, or to income levels above the reduced range.

Finally, I define the location parameter $\theta(r)$, allowing this parameter too to vary with log receipts r. The parameter θ identifies the mode of the Laplace distribution. For the symmetric Laplace distribution it is also the median, and the sample median is the maximum likelihood estimator of θ, but this is not the case when one allows for arbitrary nonsymmetry of the distribution. I therefore estimate the mode nonparametrically by year and level of current receipts using a bivariate kernel density estimator in r and growth g. I estimate the density for fifty points in the range of $r < 14$ and $-.05 < g < .1$, where the majority of mass lies. For each value of r among these fifty points I keep the level of g with greatest mass and then regress g on a quadratic function of r and use the predicted value of g for each r as the value of θ for observations with that level of base-year receipts. In practice, setting $\theta = 0$ provides nearly identical results because the estimated mode is close to 0 for all years, particularly for observations near or below the notch.

The observed conditional cdf is

count of omitted and not the location of either.

[18] 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.

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

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

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).