A Unified Welfare Analysis of Government Policies

Nathaniel Hendren, Harvard University
Ben Sprung-Keyser, Harvard University

July 2019

Abstract

We conduct a comparative welfare analysis of 133 historical policy changes over the past half-century in the United States, focusing on policies in social insurance, education and job training, taxes and cash transfers, and in-kind transfers. For each policy, we use existing causal estimates to calculate both the benefit that each policy provides its recipients (measured as their willingness to pay) and the policy's net cost, inclusive of long-term impacts on the government's budget. We divide the willingness to pay by the net cost to the government to form each policy's Marginal Value of Public Funds, or its "MVPF". Comparing MVPFs across policies provides a unified method of assessing their impact on social welfare. Our results suggest that direct investments in low-income children's health and education have historically had the highest MVPFs, on average exceeding 5. Many such policies have paid for themselves as governments recouped the cost of their initial expenditures through additional taxes collected and reduced transfers. We find large MVPFs for education and health policies amongst children of all ages, rather than observing diminishing marginal returns throughout childhood. We find smaller MVPFs for policies targeting adults, generally between 0.5 and 2. Expenditures on adults have exceeded this MVPF range in particular if they induced large spillovers on children. We relate our estimates to existing theories of optimal government policy and we discuss how the MVPF provides lessons for the design of future research.

We first and foremost thank the several hundred researchers whose empirical results form the foundation of our estimates. We are also deeply indebted to a wonderful team of research assistants: Caroline Dockes, Harris Eppsteiner, Adriano Fernandez, Jack Hoyle, Omeed Maghzian, Kate Musen, and the rest of the exceptional team of Pre-Doctoral Fellows at Opportunity Insights. We are also grateful to Raj Chetty, David Deming, Winnie van Dijk, Amy Finkelstein, John Friedman, Andrew Goodman-Bacon, Hilary Hoynes, John Eric Humphries, Larry Katz, Sarah Miller, Evan Soltas, Larry Summers, Michael Stepner, and Laura Wherry for helpful comments and suggestions, along with seminar participants at Georgetown, LSE, IFS, the University of Kentucky, the University of Minnesota, and Texas A&M. This research was funded by the National Science Foundation (#CAREER1653686 (Hendren) and #DGE1745303 (Sprung-Keyser)), the Sloan Foundation (Hendren), the Bill and Melinda Gates Foundation (Hendren), and the Chan Zuckerberg Initiative (Hendren). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

I Introduction

What government expenditures are most effective at improving social well-being? Are in-kind transfers preferable to cash transfers? Does government-provided social insurance efficiently address market failures? Should we invest more in low-income children? If so, at what age? Should they be direct investments or subsidies to parents?

A large empirical literature estimates the causal effects of historical government policies. These papers frequently conclude with a brief welfare analysis. The method of that analysis, however, often differs from paper to paper. When reporting the effects of health insurance expansions, it is common to report cost per life saved (e.g. Currie and Gruber, 1996). Studies of tax policy changes often report the implied marginal excess burden or the marginal cost of funds (e.g. summarized in Saez et al., 2012). Higher education analyses often report the cost per enrollment (e.g. Kane, 1994; Dynarski, 2000). The early childhood education literature often reports a social benefit-cost ratio (e.g. Heckman et al., 2010). These varying welfare measures make it difficult to compare policies, especially if one wishes to take a birds-eye view and perform welfare analysis across policy categories.

This paper conducts a comparative welfare analysis of 133 historical tax and expenditure policies implemented in the US over the past half-century. We focus on policies in four domains: social insurance (e.g. health, unemployment, and disability insurance), education (e.g. preschool, K-12, college, job and vocational training), taxes and cash transfers (e.g. top tax rates, Earned Income Tax Credit (EITC), Aid to Families with Dependent Children (AFDC)), and in-kind transfers (e.g. housing vouchers, food stamps). We draw upon existing analyses of the impacts of these policies to construct both the benefit that each policy provides to its recipients and the policy's net cost to the government. Benefits are captured by the willingness to pay of policy recipients. The net cost combines both initial program spending and the long-run impact of the policy on the government's budget (i.e. fiscal externalities). We then take the ratio of the benefits to net government costs to generate each policy's Marginal Value of Public Funds (MVPF).[1] Putting these components together allows us to measure each policy's "bang for the buck."[2]

The MVPF is useful because it measures the shadow price to the government of delivering

[1] See Mayshar (1990); Slemrod and Yitzhaki (1996, 2001); Kleven and Kreiner (2006) for original definitions, and Hendren (2016) for a comparison of the MVPF to alternative measures of welfare.

[2] In several cases where authors constructed their own MVPFs, we incorporate those estimates directly. Where applicable, we adjust these estimates in order to harmonize assumptions (e.g. interest rates). In cases where previous literature has conducted comprehensive cost-benefit analyses of a policy, we draw upon the components of those analyses to reformulate them into their implied MVPF.

welfare to the beneficiaries of the policy. For point of reference, a simple non-distortionary transfer from the government to an individual would have an MVPF of one. The cost to the government would be exactly equal to the individual beneficiary's willingness to pay. The MVPF can differ from this benchmark value of one if individuals value an expenditure at more or less than its resource cost. For instance, if the government provides insurance, willingness to pay may be greater than the resource costs of provision individuals if the insurance provides consumption smoothing benefits. By contrast, willingness to pay may fall below resource costs if individuals distort their behavior in order to receive higher transfers.[3] The MVPF may also deviate from the benchmark value of one if the policy induces fiscal externalities. For example, if spending a dollar on a government policy caused individuals to work less, government tax revenue might fall slightly and then the net cost of the policy would rise above \$1. By contrast, if spending that dollar caused them to get more schooling and consequently increased their income, government revenue would rise and the net cost of the policy would fall below \$1. In some cases, positive fiscal externalities may be large enough to fully offset the initial cost of the policy. In that instance, the policy has an infinite MVPF and, consequently, spending on the policy results in a Pareto improvement.[4]

More generally, comparisons of MVPFs correspond to precise statements about social welfare using the intuition of Okun's leaky bucket experiment (Okun, 1975). Given two policies, A and B, suppose $MVPF_A = 2$ and $MVPF_B = 1$. Then, one prefers more spending on policy A financed by less spending on policy B if and only if one prefers giving \$2 to policy A beneficiaries over giving \$1 to policy B beneficiaries. Whether this is desirable ultimately depends on one's own social preferences for the beneficiaries of policies A and B. MVPFs measure the feasible tradeoffs to the government – in Okun's metaphor, the "leaks" in the bucket. By measuring these shadow prices, the MVPF provides a unified method of welfare analysis that can be applied both across and within diverse policy domains.

We outline the construction of the MVPF for seven representative examples in Section 3. At a high level, our construction of willingness to pay often relies on intuition provided by the envelope

[3]The intuition here comes from the envelope theorem. Willingness to pay for a government transfer is determined by the "mechanical cost" of that transfer. Additional costs due to behavioral responses are not valued dollar for dollar.

[4]To align with terminology in existing literature, we use various terms interchangeably to refer to the same phenomenon. Any policy with a positive willingness to pay and negative net costs we define to have an infinite MVPF. Given the negative net costs, we also say that these policies "pay for themselves" or "recoup their initial costs." In the taxation literature, this is also known as a "Laffer effect." We often note that spending on policies with infinite MVPFs results in a Pareto improvement. This is because the expenditure is valued by beneficiaries and has no net cost on the government. This final claim regarding Pareto improvement formally assumes that all beneficiaries have positive willingness to pay, which is natural in many of our contexts in which the policies expanded the choice sets of all beneficiaries.

theorem. Our construction of net government costs involves calculating changes in taxes paid and transfer received, along with savings or additional costs from crowd out of other government spending. In Appendices A-F we also provide a detailed explanation of how each MVPF in our sample is calculated. As is common with any welfare analysis, the creation of our MVPFs requires various judgment calls. We, therefore, conduct an extensive set of robustness analyses, examining our assumptions about interest rates, tax rates, and forecasting methods.[5] In addition, many MVPF estimates for individual policies contain considerable sampling uncertainty. We address this by constructing category averages that pool across multiple policies and help improve the precision of our conclusions. We also test and correct for publication bias using the methods of Andrews and Kasy (Forthcoming). Given these potential sources of uncertainty, we also focus our results on broad patterns in the data, rather than conclusions about individual policies.

Our analysis is inevitably constrained by the scope of existing literature. Not all policies have been studied with the same degree of completeness. For each policy, we incorporate all affects that can reliably be translated into the MVPF, but an omitted impact could affect our welfare analysis. We therefore assess the robustness of our broad patterns to sample restrictions focused on more comprehensively studied policies. In addition, we discuss how the MVPF of each particular policy may vary with the addition (or removal) of certain effects.[6] We hope this latter analysis serves to document where future work may help refine our MVPF estimates. Throughout, we focus our primary conclusions on the broad lessons that are robust to these analyses.

Main Results Our estimates reveal a stark pattern: MVPFs vary substantially based on the age of each policy's beneficiaries. We find the highest MVPFs for direct investments in the health and education of low-income children. This includes Medicaid expansions, childhood education spending, and expenditures on college. In many cases, these policies actually pay for themselves in the long-run. Children pay back the initial cost as adults through additional tax revenue and reduced transfer payments. For example, we examine four major health insurance expansions to children over the last 50 years. We calculate an average across those policies and find that for each $1 of initial expenditure they repaid $1.78 back to the government in the long-run. In particular, we find that three of four policies fully repaid their initial costs.

We find high MVPFs for policies targeting children *throughout* childhood. We do find high

[5] We also provide a Stata do-file for each program that is available on GitHub. These programs allow researchers to easily modify the set of input assumptions into each MVPF beyond the robustness we readily provide in the paper and the Appendix.

[6] In many cases, we provide an extended discussion of these in the Online Appendix.

MVPFs for early childhood education programs, including an MVPF of roughly 44 for Perry Preschool and 12 for Abecedarian.[7] In addition, we find large MVPFs for policies targeting older children, such as historical equalizations in K-12 school financing (studied in Jackson et al. (2016)) and policies increasing college attainment. Our broad patterns contrast with the notion that opportunities for high-return investment in children decline rapidly with age (Heckman, 2006).

Our results show lower MVPFs for policies targeted to adults. Most of these MVPFs lie between 0.5 and 2. For example, we find MVPFs ranging from 0.40-1.63 for health insurance expansions to adults, 0.42-1.07 for in-kind transfers such as housing vouchers and food stamps, and 0.11-1.20 for tax credits and cash welfare programs to low-income households. These lower MVPFs reflect the fact that spending on many of these policies reduced labor earnings. This lies in contrast to our finding that many policies spending on children increased later-life earnings.

It is important to note that these differences in returns by age represent general patterns, but do not hold uniformly. There are a number of exceptions. For child policies, we find large variation in MVPFs across policies, with some estimates relatively close to 1. In particular, we find lower MVPFs for job training programs and for college subsidies that do not lead to increases in attainment. We also find lower MVPFs for transfers to disabled children and their families. This latter case illustrates that policies with lower MVPFs are not necessarily "undesirable" – they can be welfare-enhancing depending on one's social preferences. Unlike expenditures with infinite MVPFs, policies with low MVPFs involve a budgetary tradeoff that should be weighed against one's preference for redistribution.

Amongst expenditures on adults, we find relatively large MVPFs for changes in top marginal tax rates, with estimates from 1.16 to infinity. There is, however, substantial sampling uncertainty in these estimates.[8] We also find high MVPFs for spending on adults that generates spillover effects on children. For example, the provision of vouchers with counseling services to families residing in high-poverty public housing (as part of the Moving to Opportunity Experiment) helped these families move to lower poverty neighborhoods. This led to large increases in children's earnings in adulthood that generated sufficient tax revenue to pay for the program cost. Our results highlight

[7] In our baseline specifications that harmonize government revenue components across policies, we estimate that the government recoups 92% of the upfront cost of Perry Preschool and 78% of the cost of Abecedarian. Because the cost of crime impacts are often difficult to quantify, they are not included in our baseline analyses (when crime estimates are available, we we incorporate them in alternative specifications discussed in the Appendix for each policy). In this case, if one includes additional estimated effects such as the cost of crime, we estimate that Perry Preschool does pay for itself, and Abecedarian pays for 92% of the upfront cost.

[8] For example, we estimate an infinite MVPF for the 1981 reduction in the top marginal income tax rate from 70% to 50%. Our confidence interval, however, includes both 1 and infinity.

the value of further work to uncover when such spillovers are likely to occur.

Relation to Previous Theories Our MVPFs also speak to a large theoretical literature on optimal government policy. In Section IV, we discuss how to interpret our results through the lens of five major theoretical conclusions. First, we relate our results to the literature on optimal redistribution and taxation (Mirrlees 1971; 1976). The MVPFs measure the cost, or shadow price, of redistribution across individuals in society. We tend to find higher MVPFs for tax cuts to top earners as opposed to low-income households, consistent with what would be expected if a progressive planner set the tax rate.

Second, we use the logic of Atkinson and Stiglitz (1976) to evaluate the case for tax and transfers relative to in-kind expenditures (Atkinson and Stiglitz, 1976; Hylland and Zeckhauser, 1981). We highlight several cases where in-kind transfers may be more effective than cash, in particular policies with spillover effects on children. Third, we discuss the use of tagging (Akerlof, 1978) and how our results potentially bolster the case for age-dependent policies (Weinzierl, 2011). Fourth, we relate our MVPF estimates for social insurance policies to the large literature on market imperfections in insurance markets, such as adverse selection (Akerlof, 1970). We find some evidence of welfare gains from targeting such market failures, but our results suggest the most welfare-relevant market failure is the systematic under-investment in the human capital of low-income children. Finally, we discuss the case for intervention in human capital markets in light of research on optimal subsidies (Stantcheva, 2017; Bovenberg and Jacobs, 2005). We find high MVPFs for policies that promote human capital. Yet, our results suggest that policies promoting investments through the relaxation of informational and credit constraints may be more efficient than tax subsidies.

Lastly, we show our results are inconsistent with the idea that the suite of historical policy choices we observe were made by policymakers maximizing social welfare.[9] Rather, the persistently high MVPFs for policies targeting low-income children over the past 50 years are consistent with systematic inefficiencies that prevent welfare-enhancing long-run investments in children.

Implications for Future Research We conclude by providing three lessons for future research. First, we show how the MVPF framework allows us to quantify the value of such research. Because the MVPF is a shadow price, one can use a standard decision-theoretic framework to quantify

[9]We arrive at these conclusions using the logic of the inverse optimum approach (e.g. Bourguignon and Spadaro (2012)). The MVPF reflects the shadow price of redistribution, or more formally the Lagrange multiplier on the government budget constraint (see Eissa et al. (2008); Hendren (2016, 2017a) for a discussion). As such, it can be used to understand the implicit social preferences that rationalize indifference to the policy change.

the value of reducing uncertainty in our MVPF estimates. Just as a consumer would be willing to pay to learn the true value of the products he or she buys, a welfare-maximizing government should be willing to pay to reduce uncertainty in the cost of redistribution. Using this approach, we show that a welfare-maximizing government deciding whether to raise taxes in order to spend an additional $1 on the Supplemental Nutrition Assistance Program (SNAP) would be willing to pay $0.24 to make this decision using a more precise causal estimate of the long-run impact of SNAP using administrative data (as in Bailey et al., 2019) as opposed to survey data (as in Hoynes and Schanzenbach, 2009).[10] This highlights the value of expanding the access to, and use of, large administrative linked datasets for the study of long-run policy impacts on children.

Second, we show the added insights that come using the MVPF framework as opposed to traditional cost-benefit analysis.[11] As we demonstrate in Section II, the MVPF corresponds to precise statements about social welfare because it is the shadow price to the government of increasing the well-being of the beneficiaries of a policy. It turns out that our general findings would be very similar in a traditional cost-benefit framework – our conclusions do not require that one convert to the MVPF framework. That said, the MVPF leads to different conclusions in certain key instances. For example, when taxes are at the top of the Laffer curve the social benefit of reducing taxes by $1 is $2,[12] but the MVPF of that policy is infinite because the benefits to the individual are $1 and the net cost of the policy is $0. This divergence occurs because of the MVPF's focus on the government's budget constraint. More broadly, our results suggest there is value in calculating the MVPF in other settings, such as crime policy or tax enforcement, where the causal effects of the policy have clear impacts on the government's budget.

Lastly, we discuss the implications of the MVPF framework for future empirical designs. In particular, we highlight the importance of determining whether willingness to pay is positive or negative. In this paper we sought to analyze state-level welfare reforms from the 1980s and 1990s. There were 27 large-scale state-level randomized controlled trials (RCTs) analyzing welfare reform. These studies increased our understanding of the employment and revenue impacts of welfare

[10] In contrast, the value to learning the true impact of the 1986 earned income tax credit expansion for single childless adults is less than $0.01. While estimates of the impact of the EITC on single childless adults are of course still crucial for understanding longer run trends in labor force participation, they do not appear to have a meaningful effect on the welfare consequences of the policy.

[11] The edited volume from Weimer (2009) provides a discussion of cost-benefit analyses from different researchers in a range of different domains. The Washington State Institute for Public Policy (WSIPP (2018)) conducts ongoing cost-benefit analyses to assess policies relevant to state legislatures. See also Rea and Burton (2018) for an application of the WSIPP data to comparative welfare analysis.

[12] The individual is willing to pay $1 for the tax cut and the government receives a $1 benefit from increased tax revenue from the behavioral response to the tax. In traditional cost-benefit analysis, increases in government tax revenue are included in the numerator of the expression.

policy. They demonstrated that these welfare reforms had low net costs. That said, while the treated participants in these studies often received additional services such as job search assistance, these policies also cut benefits for those who did not comply with program requirements. As a result, it is unclear whether willingness to pay for these reforms was positive or negative. Despite randomizing over 100K families into 27 large-scale RCTs, we are unable to reach any reliable estimates of the MVPF of these policies. The evaluations of welfare reform may have led to more valuable information if the RCT designs had been created with a social welfare framework in mind.

Relationship to Existing Literature In constructing our MVPFs and presenting evidence for high returns to investment in low income children, we build upon a substantial line of existing research making the argument for investment in children.[13] As noted above, Heckman (2006) argues that benefits of early expenditures on children substantially exceed the costs. They also maintain that there are diminishing returns as children age. In contrast, we find the potential for high returns to government policy spending on individuals throughout childhood. Our work is also related to recent research on the long-run effect of safety net protections for children reviewed by Hoynes and Schanzenbach (2018). In light of the evidence, they conclude that "reallocation of investments over the life course to earlier periods can be efficiency-enhancing," which aligns with our conclusions. The MVPF helps to quantify these statements, and to assess the existence of potential Pareto improvements.

There are also analyses - many of which we draw upon in this paper - in which researchers have previously argued that some government expenditures largely pay for themselves. This argument is particularly prominent in discussion of early education (e.g. García et al., 2017; Heckman et al., 2010) and child healthcare expenditures (e.g. Brown et al., 2015; Wherry et al., 2018).[14] The argument also appears in the tax literature, where some have argued that reducing top marginal tax rates produces a "Laffer effect," raising total revenue.[15] Our analysis builds upon that work by evaluating policies at scale and searching for the presence of high return policies across a range of policy domains. For example, we can compare the historical evidence for the presence of Laffer

[13]For example, foreshadowing many of our conclusions, Currie (1994) writes "Although the evidence is incomplete, it suggests that in-kind programs have stronger effects on children than cash transfers, and that programs that target specific benefits directly to children have the largest positive effects."

[14]Outside the scope of this paper, some suggest certain macroeconomic policies can pay for themselves, such as fiscal expansions during deep recessions (DeLong et al., 2012). More generally, we omit many potentially relevant categories of policies, such as macroeconomic stabilization, infrastructure investment, and environmental policies, to name a few.

[15]In this sense, testing whether the MVPF of a policy change is infinite is a generalization of Werning (2007)'s proposed test for identifying local "Laffer effects" in the income tax schedule.

effects in tax cuts for adults and investments in children. We find the most robust evidence for Laffer effects for policies investing directly in children.

Roadmap The rest of this paper proceeds as follows. Section II presents the general social welfare framework that motivates the construction of the MVPF. Section III discusses the sample and presents six example constructions of the MVPF. Section IV discusses our main results and the distinction between MVPFs of policies targeting children versus adults. Section V places the MVPF estimates in the context of existing theories of optimal government policy. Section VI presents lessons for future work. Section VII concludes. As noted above, Appendices A-F provide step-by-step details for the construction of each MVPF, and all Stata do-files for the construction of each MVPF are available on GitHub.

II MVPF Framework

This section presents a general framework to measure the welfare impact of changes in government policies. The framework illustrates how the marginal value of public funds provides natural guidance on the social welfare impact of economic policies.

Consider a government seeking to measure the welfare impact of a government policy change under consideration. We define social welfare, W, by the weighted sum of individual utilities,

$$W = \sum_i \psi_i U_i$$

where U_i is individual i's utility function and ψ_i is their social welfare weight. The latter measures how much a 1 unit increase in utility corresponds to an impact on social welfare, W.[16] The utility function, U_i, measures both current and future well-being of the individual. For example, if utility were additive over time, one could nest uncertainty about future outcomes within this framework, letting $U_i = E\left[\sum_{t\geq 0} \beta^t u_{it}\right]$ where u_{it} is the individual's utility t periods from today.

Because the utility function is allowed to vary arbitrarily across individuals, it will be helpful to normalize units across individuals. To that aim, let λ_i denote individual i's marginal utility of income at the time the policy is under consideration. This is equal to the impact on individual utility of providing \$1 to that individual. Let $\eta_i = \psi_i \lambda_i$ denote the individual's social marginal

[16] For now, we do not place any assumption on these weights, and therefore they can result from any particular social welfare function. We also assume the weights do not change in response to the policy, but this is without loss of generality because we focus on small policy changes below.

utility of income at the time of the policy. The value of η_i measures the impact on social welfare, W, of an additional \$1 placed in individual i's budget today.

The government is considering a set of policy changes indexed by $j = 1, ..., J$ that change the economic environment (e.g. prices, public goods, etc.) by a small amount. We parameterize the upfront initial spending on policy j by dp_j (which can either be an increase or decrease). The net impact on social welfare of the policy is

$$\frac{dW}{dp_j} = \sum_i \psi_i \frac{dU_i}{dp_j} = \sum_i \eta_i WTP_i^j = \bar{\eta}_j \sum_i WTP_i^j \tag{1}$$

where $\sum_i WTP_i^j$ is the sum of individual's willingness to pay for policy j out of his/her own income, $WTP_i^j = \frac{dU_i}{dpj}/\lambda_i$, and $\bar{\eta}_j$ is the average social marginal utility of the beneficiaries of the policy,

$$\bar{\eta}_j = \sum_i \eta_i \frac{WTP_i^j}{\sum_i WTP_i^j}$$

with weights given by the economic incidence of the policy, $\frac{WTP_i^j}{\sum_i WTP_i^j}$. The values $\bar{\eta}_j$ measure how much social welfare increases if one were to provide an average of \$1 to the beneficiaries of policy j. Each individual is willing to pay WTP_i^j for the expansion by dp_j of policy j. Therefore, multiplying $\bar{\eta}_j$ by $\sum_i WTP_i^j$ measures the impact on social welfare of an expansion of the policy by dp_j. This means that the welfare effect depends on the impact of providing \$1 to a policy's beneficiaries, $\bar{\eta}_j$, and the beneficiaries' willingnesses to pay for the policy relative to cash, $\sum_i WTP_i^j$.

In accounting for costs, we let R denote the present discounted value of the government budget, and let $G_j = \frac{dR}{dp_j}$ denote the net impact of the policy on the government budget.[17] This net cost is inclusive both of the initial cost of the program and all other impacts of behavioral responses on the government budget. For example, if spending \$1 on preschool increases wages in the future, G_j should incorporate the impact of those increases in future tax receipts.

The Marginal Value of Public Funds (MVPF) of policy j is given by the aggregate willingness to pay, $WTP^j = \sum_j WTP_i^j$, for the policy divided by the net cost to the government, G_j:

$$MVPF_j = \frac{\sum_i WTP_i^j}{G_j} = \frac{WTP^j}{Net\ Cost} \tag{2}$$

[17] In practice, the dp_j variations that are identified in an empiricist's regressions will not, in general, correspond to budget-neutral policies. Traditional approaches would attempt to account for government spending by modifying the observed policy into a different policy that raised revenues via lump-sum taxation. This would then require the researcher to observe not the causal effect of the policy, but rather the "compensated effect" of the policy in order to identify the welfare impact. In contrast, our approach hypothetically closes the budget constraint by comparing two MVPFs: one that involves an increase in spending and another that involves a reduction in spending or increase in revenue. Hence, welfare analysis can be done with two sets of causal effects (one for the two policies under consideration) as opposed to attempting to measure the compensated effect of a policy.

The MVPF is previously defined in Mayshar (1990) where it is referred to as the marginal excess burden (MEB), in Slemrod and Yitzhaki (1996) where it is referred to as both the marginal cost of funds and/or the marginal benefit of projects, depending on the policy in question and in Kleven and Kreiner (2006) where it is referred to as the marginal cost of funds (MCPF). However, the MVPF formally differs from both the traditional definition of the marginal excess burden in Auerbach (1985); Auerbach and Hines (2002) and the marginal cost of funds in Stiglitz and Dasgupta (1971); Atkinson and Stern (1974). Because of this, Hendren (2016) defines this quantity as the MVPF to contrast it with the MEB and MCPF.

Combining equations (1) and (2), the impact on social welfare per dollar of government expenditure on policy j is

$$\frac{\frac{dW_t}{dp_j}}{\frac{dR}{dp_j}} = \bar{\eta}_j MVPF_j$$

Given the MVPF for any two policy changes, one can construct hypothetical budget-neutral policy change. For example, consider increasing spending on policy 1 by a net amount G_1, financed by reducing spending (or increasing revenue) from policy 2 by a net amount of G_2. Pursuing this combined policy, dp, increases social welfare if and only if

$$\bar{\eta}_1 MVPF_1 > \bar{\eta}_2 MVPF_2. \tag{3}$$

Welfare increases if and only if the welfare gains from increasing spending on policy 1, $\bar{\eta}_1 MVPF_1$, exceed the welfare loss from reducing spending on policy 2, $\bar{\eta}_2 MVPF_2$. The MVPFs of the two policies characterize the cost of moving welfare between the two groups of beneficiaries. One prefers the policy if and only if $\frac{\bar{\eta}_1}{\bar{\eta}_2} > \frac{MVPF_2}{MVPF_1}$. If $MVPF_1 = 1$ and $MVPF_2 = 2$, then an individual prefers spending on policy 1 financed by policy 2 if and only if providing \$1 to beneficiaries of policy 1 is valued more than providing \$2 to beneficiaries of policy 2.

As this example illustrates, welfare statements that compare policies generally require comparisons of their MVPFs. The MVPFs allows the researcher to form hypothetical budget neutral policies and assess their welfare implications using equation (3). To reduce the role of social preferences in driving conclusions, one can compare policies with the same beneficiary group. In this case, one would expect that $\bar{\eta}_1 \approx \bar{\eta}_2$ so that comparisons of the $MVPFs$ correspond to statements about social welfare. For example, Hendren (2017a) suggests comparing the MVPF of a particular policy to the MVPF of a tax cut with similar distributional incidence. More generally, one can also compare different redistributive policies such as food stamps and housing vouchers amongst each other to evaluate the most effective method of redistribution.

In some cases, one does not need to compare an MVPF to another policy to reach a welfare conclusion. This occurs when the MVPF is infinite. Mathematically, this happens when a policy has positive willingness to pay by its beneficiaries and the behavioral response to the policy generates fiscal externalities that are sufficient to cover the cost of the program, $G_j < 0$. The canonical example of such a case is lowering taxes when they are beyond the peak of the Laffer curve. In this case, lowering taxes increases government revenue and so these policies represent a Pareto improvement for any positive welfare weights assigned to the recipients.[18] More generally, the MVPF framework facilitates a search for Laffer effects of other policies, such as investment in children.

Comparison to Social Cost-Benefit Analysis A more common approach to welfare analyses in previous literature is a social cost-benefit analysis (Heckman et al., 2010). To form this, let WTP^j denote the aggregate willingness to pay for policy j. Let C_j denote the initial program outlays of the policy (absent the impact of any behavioral responses on net costs), and let $FE_j = G_j - C_j$ denote the impact of the behavioral responses to the policy on the government budget. The benefit cost ratio is then given by:

$$BCR_j = \frac{\text{Social Benefits}}{\text{Social Costs}} = \frac{WTP^j + FE_j}{(1+\phi)\, C_j} \tag{4}$$

The numerator is the net social benefits of the policy, inclusive of the benefits that accrue back to the government. The initial program outlays in the denominator are often multiplied by $1 + \phi$, where ϕ is the marginal deadweight loss of raising government revenue. This is thought to translate the upfront costs into social costs by accounting for the welfare impact of an implicit tax policy that raises the needed funds. Often, ϕ is taken to be 0.3 or 0.5 (Heckman et al., 2010). Policies are then deemed to pass the cost-benefit test if the BCR exceeds 1.

In contrast to the BCR, the MVPF is given by $MVPF_j = \frac{WTP^j}{C_j + FE_j}$. It differs in two primary ways. First, the impact of behavioral responses on the government budget are counted in the denominator, not the numerator. For example, consider a tax cut of \$1 for which the behavioral response increases tax revenue by \$1. In this case, the policy perfectly pays for itself, and so the MVPF is infinite. Expenditures on the policy represent a Pareto improvement. In a BCR

[18] In practice an expenditure policy may have been combined with a separate tax policy to raise revenue at the time the policy is implemented. In this case, the combined expenditure and tax policy would not deliver a Pareto improvement, as some current taxpayers would be made worse off. However, the infinite MVPF corresponds to a case where the government need not raise revenue to implement a policy that does not cost money in the long-run. The government could have borrowed against the future returns on the policy and generated a Pareto improvement.

framework, however, that \$1 in increased tax revenue is considered social benefit and counted in the numerator. That leaves a BCR estimate of $2/(1+\phi)$. This illustrates why the BCR may be a particularly misleading guide to optimal policy when policies have strong impacts on the government budget. We found a policy with a BCR of $2/(1+\phi)$ that was a Pareto improvement, but we could find a different policy with a BCR above 2 that does not deliver a Pareto improvement. For example, if we compare this hypothetical tax cut to government provided insurance with willingness to pay of \$2 for each \$1 of insurance, the traditional cost-benefit framework cannot distinguish between these policies.

Second, when constructing the MVPF one does not adjust for the "deadweight cost of taxation" based on particular assumed method of government finance. Rather, the MVPF is the shadow price of the policy in question: it measures how much welfare is delivered to beneficiaries per dollar of government expenditure. One simply evaluates the MVPF of a given policy relative to the MVPF of other policies. This allows for comparisons across broad sets of policies and allows the researcher to think through the library of feasible levers available to the government. In contrast to the cost-benefit framework, this approach reinforces the idea that incidence matters: a policy that provides benefits to the poor cannot be readily compared to the raising of revenue on the rich without thinking about Okun's bucket and the social welfare weights placed on the beneficiaries (i.e. the values of $\bar{\eta}_j$ for the policies).

Despite our advocacy for the value of the MVPF over a traditional cost-benefit analysis, it is perhaps reassuring to note that, in most cases, these two approaches generate similar conclusions. So, although we argue that the MVPF is more appropriate for measuring welfare, and consequently, more informative in cases where these two welfare measures diverge, the broad pattern of our results remain the same under either framework.

III Calculating MVPFs: Examples

We estimate the MVPF for 133 policies spanning social insurance (e.g. health, unemployment, and disability insurance), education (e.g. preschool, K-12, college, job and vocational training), taxes and cash transfers (e.g. top tax rates, EITC, AFDC), and in-kind transfers (e.g. housing vouchers, food stamps). Our focus here is on policies, rather than papers. That means in many cases we combine estimates from multiple different papers, fitting together the puzzle pieces to build the full

picture.[19]

We form a sample of policies in each domain by drawing upon survey and summary articles from each field. We then supplement this initial set of estimates with recent work in each area not captured in the survey or summary articles. We restrict our attention to policies in which there is an experimental or quasi-experimental identification strategy used to estimate the policy's impact.[20] Formally, such papers identify causal effects using variations dp_j in the economic environment. We form our baseline sample with policies where one observes impacts of the policy that are sufficient to form a reasonably comprehensive view of both the WTP and net cost of the policy. We discuss in Appendices A-F the standard for policy inclusion in each of our categories and the set of causal effects used in each case. Since this process involves judgment calls, we also assess robustness of our conclusions to an expanded sample (e.g. that expands the set of identification and forecasting methods) and a more restricted sample (e.g. that requires direct observation of causal impacts on income).

Table 1 lists the set of policies studied, along with the empirical papers used to form each policy's MVPF. Column 9 denotes the set of papers used to construct the MVPF. In many cases, we draw upon multiple papers to form a single MVPF. For example, some papers might estimate the impact of the policy on adults, while other papers focus on longer-run impacts on children.

In this section, we illustrate the construction of these estimates using seven examples spanning each of the domains we consider. We attempt here to provide a diverse set of examples to demonstrate the range of approaches used to create our estimates. Appendices A-F provides a detailed step-by-step discussion of the construction of each MVPF. In Section IV.C, we assess robustness of our primary conclusions to alternative assumptions (e.g. different interest rates and tax rate imputations) and alternative samples. We also provide a Stata do-file for each program that is available on GitHub. These allow researchers to investigate in detail how each MVPF estimate was created and utilize alternative assumptions in future work.

III.A Admission to Florida International University

We begin by constructing the MVPF of admitting an additional student into Florida International University (FIU). This example illustrates the construction of the MVPF for a policy targeting

[19] If multiple papers analyze the same causal effect, we generally focus on the most recent published estimates unless otherwise noted. We provide a detailed discussion of the alternative specifications in the Appendix.

[20] We exclude purely cross-sectional identification using controls for observables in our baseline sample. We do not impose our own filter on the quality or validity of these empirical designs, but focus our sample on papers discussed in leading surveys of each of these literatures.

youth with impacts on later-life earnings. We use similar methods for other child policies.

We draw upon the work of Zimmerman (2014). He uses an RD design at the school's academic performance cutoff for applicants to measure the impact of FIU admission on state university system enrollment and medium-term earnings outcomes. We translate his estimates into an MVPF, incorporating the net cost of the policy and the beneficiaries' willingness to pay. Throughout, we construct confidence intervals for our estimates using a semi-parametric bootstrap procedure discussed in detail in Appendix H.[21]

Costs Figure IA shows how we calculate the net cost of FIU admission. We start with initial costs of $11,403, which represents the state university system's educational expenditures on each marginal admit to FIU.[22] Students pay some fraction of those educational expenses, and so we subtract off $3,184 to account for private student contributions. Next, we account for the fact that some new admits would have attended a state community college if they had not enrolled in FIU. We subtract off $5,601, Zimmerman's estimate of the amount the government would have paid to support their education at those community colleges. Taken together, that leaves us with a upfront government cost of $2,617 per admitted student.

The remaining cost considerations all stem from earnings changes caused by FIU admission. Zimmerman (2014) calculates that in the first seven years after admission, earnings fall by $10,942.[23] We use estimates from the Congressional Budget Office to estimate that the tax and transfer rate on these earnings is 18.6%. This suggests the earnings change reduces government revenue by $2,035.[24] Next, Zimmerman (2014) estimates that FIU admission causes earnings to rise by $36,369 in years 8-14. Once again, we apply a tax and transfer rate and we determine the government revenue

[21] In particular, we conservatively account for correlations across estimates within a given policy, and we develop a method to adjust for the uncertainty in the denominator (with many thanks to conversations with Isaiah Andrews). We provide both the intuition for the approach and we provide Monte Carlo simulations with appropriate coverage. In fact, the coverage is sometimes overly conservative, especially when costs approach zero.

[22] Zimmerman (2014) calculates these costs using the data on educational expenditures from the Delta Cost Project (American Institutes for Research, 2017). We adopt this approach in calculating costs for other college policies analyzed in our sample. Appendix B explains the details of our approach.

[23] All earnings changes are discounted back to the time of the initial expenditure using a 3% discount rate. We toggle these discount rates in our robustness discussion in Section IV.C. We also use CPI-U-RS when we need to deflate from nominal dollar values to real ones.

[24] To be conservative, we exclude payroll taxes because individuals may benefit from a portion of these contributions. More detail on our calculations can be found in the Appendix G. The tax and transfer rate includes federal and state income taxes along with food stamps; but excludes housing vouchers and other welfare programs. We use the income-specific rate from the 2016 CBO estimates, and we apply this rate uniformly across years for simplicity. With more reliable historical information on marginal tax and transfer rates across the income distribution, one could perform the analysis separately by year. We are, however, not aware of any comprehensive historical source on the distribution of those rates. For this reason, we take the simpler approach of using a consistent 2016 tax and transfer rate and then assessing the robustness of all our results to alternative rate assumptions. We present robustness to alternative tax and transfer rate assumptions in Section IV.C.

rises by $7,274. At this point our net costs are $-2,622, as shown in Figure IB. This suggests the expenditure has paid for itself within 14 years of the initial outlay.

Finally, Zimmerman's earnings data extend fourteen years, but we can extrapolate from the observed effects to estimate earnings changes over the full life-cycle. Appendix I describes this procedure in detail, and Appendix Figure I provides a graphical illustration of the approach. We use ACS data to estimate life-cycle earnings trajectories and then map the control group in Zimmerman (2014) onto those trajectories. In particular, we observe an average earnings for the control group of $28,964, which we estimate to be 113% of mean earnings for this cohort in the ACS. In contrast, the treated group earns $6,372 more during these ages, or 22% more than the control group. We assume that the control group earnings remain constant as a fraction of average ACS earnings throughout the life cycle. We also assume that the percentage earnings increase for the treatment group also remains constant throughout the life cycle. These assumptions mean that we assume the trajectories for the treatment and control groups differ by a constant percentage throughout the life cycle.[25] This yields an estimated discounted earnings increase of $117,330 through age 65. We subsequently calculate that the associated fiscal externality reduces government costs by $21,823. When combined with our previous cost components, we find that each marginal FIU admission has a net cost of $-24,445. The expenditure pays for itself.

Willingness to Pay Having established that the initial costs of increasing admission at FIU leads to a long-run net savings to the government, the policy has an infinite MVPF as long as $WTP > 0$. That said, constructing a measure of willingness to pay remains useful in making our confidence intervals and evaluating alternate specifications. Broadly, we take two approaches to measuring willingness to pay: a baseline estimate and a conservative approach. The components of these calculations are illustrated in Figure IC.

Throughout, our approaches to estimating WTP rely heavily on the logic of the envelope theorem and revealed preference. For the baseline estimate, we assume that increases in income amongst the college educated stem from returns to human capital, not from higher levels of effort.[26] In this case, the envelope theorem implies that we can form an estimate of WTP using the policy's impact on net income after taxes and other expenses. We begin by noting that those who are admitted to

[25] While this is a strong assumption, we show in the robustness analysis that our results are actually not very sensitive to the method we use to construct these forecasts. For example, we conduct a conservative forecast that assumes zero income growth over the life cycle. This yields similar results (See Figure VIC).

[26] If higher returns were the result of higher levels of effort, that would require an adjustment for the disutility of labor.

FIU have an increase in private costs associated with additional tuition and fee payments at the four-year school. This leads to a negative WTP component of $2,851. Next, the earnings fall in the first seven years after admission leads to a further negative WTP of $8,907. The earnings gains in years 8-14 yield a positive WTP of $29,095. Projecting through the rest of the life cycle yields an additional WTP of $95,507. Combined, this yields a total willingness to pay of $112,844.

Our second approach forms a conservative measure of WTP that does not incorporate any willingness to pay from rising net incomes. Instead, we allow for the idea that earnings gains may have been offset with an increase in labor effort that decreases utility. In this case, we note that individuals on the margin choose to pay $2,851 more in tuition in order to enroll in FIU. This means that by revealed preference they are WTP at least $2,851 to be admitted. This approach is conservative because those who enrolled certainly may have been willing to pay more for admission than their marginal tuition charges.

III.B Medicaid Expansion to Pregnant Women and Infants

Next, we consider a Medicaid expansion to pregnant women and children in the US that occurred across states between 1979-1992. This example illustrates a case where we construct the MVPF using examples from several papers using the same identification strategy but focusing on different outcomes.

We construct our MVPF using several different analyses of these reforms, each of which use the differential timing of the reforms across states to measure their impacts. Currie and Gruber (1996) document a significant increase in health insurance coverage for pregnant women, along with a corresponding reduction in infant mortality and low birth weight. Cutler and Gruber (1996) find significant crowd-out of private insurance policies. Dave et al. (2015) find reductions in labor supply of eligible women. Miller and Wherry (2018) find positive impacts on children's future earnings and health for those whose parents obtained Medicaid eligibility. We translate these estimates into their implied MVPF, beginning with costs and then turning to willingness to pay.

Costs The bar chart in Figure IIA illustrates the translation of estimates from the literature into their implied costs to the government. Currie and Gruber (1996) estimate the cost of insuring an additional pregnant woman through the Medicaid expansion was $3,473.[27] In addition to the direct Medicaid costs, Dave et al. (2015) estimate Medicaid eligibility leads to a 21.9% reduction in female

[27]For consistency across papers analyzing the reform, we deflate all numbers to 2012 USD using the CPI-U-RS; as a result, they differ slightly from reported figures in each paper.

labor force participation, which corresponds to an earnings impact of roughly $2,834. We estimate these individuals face a tax-and-transfer rate of 19.9% from the CBO using our procedure discussed in Appendix G. This means the earnings impact implies an additional cost to the government of $564 per eligible child.

Turning to the impacts on children, Miller and Wherry (2018) estimate that a one percentage point increase in parental eligibility leads to a reduction in future hospitalizations of 0.237% when children are 19 to 32 years old. With a 3% discount rate, this implies a government savings on Medicaid and uncompensated care of $868 over the 14 year period from ages 19 to 32.[28] Miller and Wherry (2018) also find a 3.5% increase in college attendance and an 11.6% increase in earnings for children made eligible. On the one hand, to the extent to which the government subsidizes college expenses, increased enrollment raises government costs. We estimate that effect to be $371. On the other hand, the increase in earnings when children are 23-36 years old leads to an increase in government revenue of $3,909. By the time children are 36 years old, the estimates suggest the policy has paid for itself.

As with the example in Section III.A above, we forecast these earnings gains to age 65 by assuming the percentage impact on earnings remains constant throughout the life cycle. This suggests the government recoups an additional $6,114 in tax revenue over this period, for a total of $10,023. The upfront cost of $3,473 led to a long-run net government surplus of $7,014 (95% CI of [1,178, 12,971]).

Before moving on to discussing the details of willingness to pay, it is worth noting that the MVPF of this expenditure has already been determined. In order for a policy to have a infinite MVPF, net costs must be negative and willingness to pay must be any positive value. The policy evaluated here expanded healthcare opportunities to parents and children, so it is safe to assume willingness to pay is positive. In fact, if the policy did not make anyone worse off, then these expenditures resulted in a Pareto improvement.

WTP While the baseline MVPF estimate is infinite, we calculate willingness to pay for use in constructing confidence intervals and evaluating alternate specifications where costs are positive. Here, we briefly summarize this construction, which consists of three components. (Step-by-step details of this calculation can be found in Appendix D.)

First, Cutler and Gruber (1996) document that half of the increase in Medicaid actually crowded

[28] We forecast to age 65 by assuming a constant dollar savings and discounting by 3%, which implies $530 of total savings as shown in Figure II.

out private coverage. Assuming the public and private costs of insurance were roughly similar, this finding implies that beneficiaries no longer had to pay roughly $1,737 in health insurance costs. This means that WTP is at least $1,737. Second, Currie and Gruber (1996) estimate a causal effect of the Medicaid expansion on infant mortality. We assume parents have a willingness to pay out of their own income of $1M to avoid an infant death (and assess robustness to alternative specifications).[29] Third, we consider the WTP by the children for improved labor market prospects in adulthood. To do so, we assume that the increase in earnings documented by Miller and Wherry (2018) reflects an expansion of labor market opportunities and not an increase in costly labor effort. This means that the children should be willing to pay the increase in their net income after private expenses that results from increased educational attainment. The increase in after-tax income is $17,728 for the observed 14 year age range (23-36) in Miller and Wherry (2018) and an additional $26,236 in the subsequent years. Subtracting the cost of college expenses reduces this by $111 for a net WTP of $48,353.

We also provide a conservative WTP estimate using solely the transfer value of the insurance of $1,737. This would be valid if the increase in after-tax earnings came at the expense of increased effort as opposed to increased opportunities. To be sure, the difference between the conservative and baseline WTP estimate is quite large. As we discuss below, our primary conclusions remain valid.

III.C Introduction of Food Stamps

Third, we construct an MVPF for the impact of the introduction of the Food Stamp Program, today known as the Supplemental Nutritional Assistance Program (SNAP). This example illustrates how we incorporate potential spillovers of adult-targeted policies onto children.

The Food Stamp Program provides in-kind transfers to low-income families that can be used on food. Its introduction in the 1970s was staggered across counties in the US.[30] Hoynes and Schanzenbach (2012) exploit this variation to analyze its impact on labor income and welfare participation of adult beneficiaries; Almond et al. (2011) study its impact on birth outcomes. Bailey et al. (2019) use the same variation to study its impacts on the adult earnings of children

[29] Note we should think of this as a "private" not a "social" willingness to pay. It assumes that parents are willing to pay $10,000 out of their own pocket to have a 1% reduction in infant mortality. It is important to note that society may well be willing to pay more than $1M. In the language of the social welfare function, this suggests the population has a high social marginal utility of income, η_j.

[30] This variation was initially studied by Currie and Moretti (2006) in California and extended nationally by Almond et al. (2011), Hoynes and Schanzenbach (2012), Hoynes et al. (2016) and Bailey et al. (2019).

whose parents received food stamps.

Costs The first component of our total costs is the average yearly benefit from food stamp enrollment, equal to $2,904. To this, we add the fiscal externality resulting from the impacts on both adults and children. For adults, Hoynes and Schanzenbach (2012) document large yet imprecise reductions in earnings of $3650 that imply a fiscal externality of $471 from reductions in tax revenue – roughly $0.16 per $1 of food stamps provided. For children, Bailey et al. (2019) find increases in earnings in adulthood corresponding to 7.1% for 6 full years of childhood exposure to food stamps between the ages of ages 0-5. In Appendix E, we show this corresponds to an estimated increase in tax revenue of $0.24 per $1 of food stamps for every family with a child aged 0-5. We then multiply this by 0.35, the fraction of SNAP benefits received by households with children age 0-5. We subsequently multiply by 1.32, the average number of children in these households. This suggests that for each $1 in food stamp spending the resulting impacts on children increase in government revenue by $0.11.[31] Taken together these estimates imply that every $1 of spending on food stamps costs $1.05.[32]

WTP We provide a willingness to pay from three components. First, the envelope theorem suggests that individuals are willing to pay for the mechanical cost of SNAP benefits, which we estimate to be $1,809. We arrive at this number by taking the $3650 increase in earnings and noting that SNAP benefits decline with earnings at a 30% phase out rate. This means that $1095 of the food stamp cost is the result of a cost increase from behavioral responses. Consequently, our point estimate suggests individuals value $0.62 for each $1 spent by the government on food stamps.[33] Second, we incorporate the WTP for reductions in infant mortality and increases in longevity amongst their children. As in the case of Medicaid in Section III.B, we assume this is given by the reduction in child mortality multiplied by a VSL of $1M (2012USD). We add to that value the number of years of increased longevity multiplied by a QALY of $20k (2012 USD). This

[31] We assume no impact on children at older ages, but clearly such effects could alter the MVPF. In Section V.C, we discuss the implications for a policy targeted to families with children aged 0-5; this leads to a larger MVPF.

[32] Our costs estimates here are constrained by the set of observed outcomes that we can reliably translate into impacts on the government budget. For example, Hoynes et al. (2016) report that the introduction of food stamps was associated with a reduction in adult metabolic syndromes. While our earnings estimates likely capture the impact of those health changes on labor supply, we lack a reliable way to measure the impact of those health changes on healthcare utilization. Future work documenting long-run health impacts that reduce (increase) government spending on medical care could lead to a higher (lower) MVPF than we estimate here.

[33] It is also worth noting that this willingness to pay is nearly identical to the value we would receive if we did not apply the envelope theorem in this context, but rather used estimates from Whitmore (2002) suggesting food stamps have a trade value of at least 65%. For our "conservative" willingness to pay specification, we make both the envelope theorem and trade value modifications and find that the MVPF falls to 0.39.

leads to an additional WTP of $0.02. Lastly, we incorporate an additional willingness to pay due to increases in after tax income amongst those who received food stamps as children. These estimates of after tax income are based on the earnings gains we calculate above. Combining costs with willingness to pay creates an MVPF of 1.04 (95% CI of [-0.97, ∞]).[34]

It is important to note in this case that statistical uncertainty in these estimates is quite high. The combination of substantial earnings reductions amongst parent and large earnings gains amongst children mean that we cannot reject MVPFs of 0 or ∞. We will return to this uncertainty in more detail in Section VI.A when we discuss the value of additional research or data access in reducing sampling uncertainty.

III.D Paycheck Plus in New York City

Fourth, we measure the MVPF of the Paycheck Plus program. This construction illustrates how we create the MVPF from randomized controlled trials (RCTs). It also provides guidance on the ideal set of measures future researchers could construct to more directly estimate the MVPF associated with RCTs.

The Paycheck Plus program is modeled after the Earned Income Tax Credit (EITC). The EITC provides income subsidies to low-income workers that are intended to encourage employment. If workers face high marginal tax rates due to the benefit schedule for means tested transfers such as food stamps, the EITC may offset those high rates. While the EITC generally targets adults with children, the Paycheck Plus in New York City program conducted a randomized controlled trial to evaluate the provision of EITC-like benefits to single adults without dependents – a group not traditionally eligible for significant EITC benefits. The credit is worth up to $2,000 per year and is available over three years (2014-2016 fiscal tax years with bonuses paying out in 2015-2017).

Miller et al. (2017) estimate the impact of the policy on income, employment, and after-tax income for the first two years of the policy, which we translate here into their implied MVPF.[35] We begin with costs.

[34] In Appendix E we also explore several alternate specifications and find these produce only small changes to the MVPF. For example, we assume a higher VSL of $9M and a QALY of $180k and find an MVPF of 1.15. We incorporate the impact of reduced incarceration based on effects estimated in Bailey et al. (2019) costs of incarceration from Heckman et al. (2010). We find that the MVPF rises from 1.04 to 1.07.

[35] As discussed in Appendix D, the current set of results from the third year do not include sufficient information to form the MVPF in as precise of a manner as we do here; but we note how imposing a reasonable additional assumption suggests the third year impacts lead to a very similar MVPF also near 1.

Cost The cost of the policy is the observed causal effect of the policy on the government budget.[36] To measure the costs, let T^j denote the tax schedule faced by the control ($j = 0$) and treatment ($j = 1$) group. And, let y_i^j denote individual $i's$ earnings if they face the $j = 0, 1$ tax/transfer schedule. The cost is then given by:

$$Cost = E\left[T^0\left(y_i^0\right)\right] - E\left[T^1\left(y_i^1\right)\right] \tag{5}$$

Because Paycheck Plus is an RCT, we compute equation (5) using the difference in tax and transfer revenue obtained by the government. In 2014, the causal impact on government costs was \$621; in 2015, this cost was \$453. Combining these values, the cost is \$1,074.

WTP We use the envelope theorem to estimate the WTP for Paycheck Plus. In 2014, the average bonus paid is \$1,399 amongst those who take it up, and 45.9% of people do so. The envelope theorem suggests participants do not value the full \$1,399 subsidy dollar for dollar. This is because part of this cost reflects the impact of behavioral responses. To first order, those who entered the labor force in order to obtain the transfer are indifferent between working and not working. Miller et al. (2017) find a causal effect of the program on the extensive margin labor supply of 0.9%. Absent behavioral responses, this implies that 45% of the sample, as opposed 45.9%, would have received the transfer had they not changed their behavior. Consequently, 98% of the transfer (45/45.9) is valued by the beneficiaries, which implies a WTP of \$630 for the transfers in 2014.[37] Repeating this calculation using the data from 2015 yields a WTP of \$441. This suggests a 2-year WTP of \$1,071. The estimated WTP of \$1,071 combined with the net cost of \$1,074 implies an MVPF of 0.997 (which rounds to 1 in Table 2).

One can also construct an MVPF separately using the 2014 or 2015 transfers and responses. This yields similar MVPFs of 1.014 and 0.973. This dynamic similarity will be a recurring theme

[36] In the context of an RCT, our approach measures the welfare impact of randomly assigning additional people to the treatment as opposed to the control group. As a result, one can use the reduced form results to form our welfare analysis (i.e. one need not separately isolate a LATE/TOT/etc.). The denominator is the causal effect of this assignment on the budget and the numerator is the aggregate WTP by members of the control group to be in the treatment group. As a result, whether our welfare analysis can be externally generalized to a different policy with different take-up of benefits would depend on how its treatment effects vary across the population.

[37] This calculation assumes no intensive margin responses. If one observed the micro-data from the RCT, one could allow for intensive margin responses. To first order, the WTP is the mechanical change in the tax schedule (i.e. replacing T^0 with T^1) holding behavior fixed for each individual at y_i^0:

$$WTP = E\left[T^0\left(y_i^0\right) - T^1\left(y_i^0\right)\right] \tag{6}$$

This means the ideal method of calculating WTP is to feed the distribution of control group earnings into both the control and treatment group tax schedule. In practice, this number is rarely reported, but future work conducting welfare analyses of RCTs can directly construct this measure.

amongst transfer programs to adults. It means that a static model of the labor market distortions provides a reasonable approximation to measuring the MVPF for these policies. Every $1 the government spends in transfers leads to a benefit of roughly $1.[38]

III.E Job Corps

Next, we construct the MVPF for a randomized controlled trial of Job Corps, one of the largest vocational education program in the United States. This example illustrates how not all attempts to increase children's human capital and earnings have high MVPFs.

Established in 1964, Job Cops is administered by the U.S. Department of Labor and provides job training and other services to at-risk youth between the ages of 16 and 24 via a network of centers run by local public and private agencies (Schochet et al., 2008). Between 1994 and 1996, the National Job Corps study randomized 80,000 eligible applicants into the program. We form an MVPF for this RCT using the recent work of Schochet (2018) who links the original RCT to tax data; we supplement this analysis with the earlier cost-benefit analysis of Schochet et al. (2006).

Cost Schochet et al. (2006) estimates that the upfront programmatic cost per recipient is $16,158. Schochet (2018) then estimates the earnings impact of the program over the course of 20 years and finds minimal effects. In particular, they find that the program increases the present discounted value of participant earnings by $121 using a 3% discount rate. We estimate this corresponds to an increase in tax and transfer revenue of $52.[39] To these, we add the value of the products produced by the Job Corps participants, which Schochet et al. (2008) estimates to be $220. Summing, this implies a net cost of the program over 20 years of $15,886. Given the small impacts on earnings, we use this 20 year observed period as our baseline estimate. In Appendix C.VII, we show that if one extrapolates these earnings affects to age 65, the net cost of the program would fall to $15,832 due to a small subsequent earnings gain.

[38] If the provision of work subsidies today leads to an increase in labor earnings and thus tax revenue after the earnings subsidies have ended, then the MVPF would be higher. We discuss these forecasts and their implied MVPFs in Appendix F. To ensure our conclusions are not biased by including policies for adults that do not have long-run follow-ups, in Section IV.C we conduct robustness of all our analysis to policies where long-run follow-ups have been measured.

[39] As discussed in Appendix C.VII, we form this estimate by summing the observed increase in tax revenue for years 6-20 in administrative data from Schochet (2018) combined with an application of the CBO tax rate to the earnings impacts for the first five years. We note that a fiscal externality of $52 in this case corresponds to a high implicit marginal tax rate. This is driven by a low tax rate on initial earnings declines and a comparatively higher tax rate on subsequently small earnings gains.

WTP Following our approach for other policies that have the potential to increase human capital, our baseline measure of willingness to pay consists of the impact of the policy on after-tax income. This is given by the \$69 increase in after-tax earned income plus the \$2,314 component of the programmatic cost that is a transfer to participants to pay for food and clothing while participating in the program. Summing, this yields a WTP of \$2,383. Dividing by the government cost of \$15,886 yields an MVPF of 0.15. If one extrapolates the earnings affects to age 65, the resulting MVPF is 0.18.[40]

III.F Unemployment Insurance

Next we consider the MVPF of expenditures on unemployment insurance (UI). There is a large literature studying optimal UI policies, often focused on estimating the "Baily-Chetty" condition (Baily, 1978; Chetty, 2006). The Baily-Chetty condition asks whether individuals are willing to pay the net cost of additional unemployment insurance. The WTP for \$1 of additional UI benefits is often measured as $1 + \gamma \frac{\Delta c}{c}$, where γ is a coefficient of risk aversion and $\frac{\Delta c}{c}$ is the impact of unemployment on consumption.[41] The cost of \$1 of UI can be written as $1 + FE$, where the fiscal externality from providing \$1 of additional unemployment insurance is measured from the impact of additional UI on unemployment duration.

If individuals were required to pay the cost of their additional unemployment insurance benefits, then additional UI would increase (decrease) welfare in cases where $\gamma \frac{\Delta c}{c}$ is greater (less) than FE. But in practice, beneficiaries of additional UI benefits are not necessarily the ones who pay for the cost of those benefits. Therefore, we use these same components to estimate the MVPF of UI:

$$MVPF^{UI} = \frac{1 + \gamma \frac{\Delta c}{c}}{1 + FE} \tag{7}$$

While the Baily-Chetty condition generally treats UI as solely a social insurance policy, the MVPF places UI in the broader context of policies with distributional incidence. This means that additional spending on UI benefits may be desirable not solely because of its correction of market failures, but also because it could be a more efficient method of redistribution.

[40] Our analysis here focuses on the MVPF of the entire treatment group. However, it is worth noting that Schochet (2018) find larger effects for the subsample of age 20-24 participants, including a 2.4 percentage point reduction in disability insurance receipt and a roughly \$500 per year increase in earnings. To see how this could lead to a different MVPF, we can first take a back-of-the-envelope calculation of a PDV of lifetime disability insurance receipt of roughly \$200K consistent with Von Wachter et al. (2011). This implies a cost savings of $4,800$, Second, we note that the \$500 per year impact on earnings corresponds to a PDV increase in earnings of \$12.8K. Applying an approximate 20% tax and transfer rate implies an increase in WTP by \$10.2K and an increase in tax revenue of 2.6K. This implies a net cost of roughly \$8,600, which implies an MVPF of 1.18.

[41] This is derived as a first order approximation from a Taylor expansion of the utility function assuming consumption is the only argument affecting the marginal utility of income.

We build our sample of variation in unemployment insurance policies using the survey article of Schmieder and Von Wachter (2016). They survey the literature on two types of policy changes to unemployment insurance: (i) changes in the size of benefits and (ii) changes in the duration of benefit availability. Conveniently, Schmieder and Von Wachter (2016) provide estimates of both the WTP ($\gamma\frac{\Delta c}{c}$) and cost (FE) from these policies for 10 benefit changes and 3 duration changes.[42]

Costs To measure the costs, $1 + FE$, we take the estimates of the costs from Schmieder and Von Wachter (2016) directly without major changes. The FE measures the impact of expanding UI benefits on unemployment duration. This has two fiscal externality components that we incorporate: first, it increases UI benefit spending; second, government tax revenue declines due to foregone incomes while unemployed.[43]

WTP To measure WTP, $1+\gamma\frac{\Delta c}{c}$, we use estimates from Gruber (1997) reported in Schmieder and Von Wachter (2016). We make only two two minor modifications to the Gruber (1997) consumption changes, $\frac{\Delta c}{c}$. First, we adjust the consumption change estimates upward using results in Hendren (2017b) to account for the fact that consumption appears to diverge for the unemployed in the 1-2 years prior to the unemployment spell. This increases the estimated value of $\frac{\Delta c}{c}$ by 24.5%. Second, for the case of benefit extensions, we adjust the estimates using results from Ganong and Noel (2019) who measure the dynamic path of consumption throughout an unemployment spell and into benefit exhaustion. We incorporate this because the relevant consumption difference is between consumption during employment and the level of consumption at the exhaustion of benefits, not the average consumption during the spell. This change further increases the WTP for \$1 of benefit extensions by \$0.13.

Broadly, the results suggest that the WTP for UI is roughly similar or slightly below its fiscal externalities. Spanning across all the policy variations considered, we find MVPFs for UI of 0.43-

[42] We take the estimates of the FE from Schmieder and Von Wachter (2016) directly. The main distinction in their calculation relative to our methods in other programs is that they use a full "labor wedge" for the tax revenue implications, which includes payroll taxes. The MVPFs are similar but slightly higher if one instead used our method of assigning tax rates.

[43] In Appendix B, we consider the potential implications of spillover effects of additional UI on reductions in disability insurance (DI). Evidence from Lindner (2016) and Mueller et al. (2016) suggest that expanded UI benefits reduces DI claiming. While this literature has identified the impact of UI on DI claiming in the short run, the long-run size of the spillover is unclear. This depends on whether the marginal individual prevented from DI claims in response to a UI expansion would have never claimed DI, or whether the UI expansion simply changed the timing of their eventual DI claim. Our baseline FE calculation assumes no spillovers from DI (i.e. assumes that the findings in Lindner (2016) and Mueller et al. (2016) reflect re-timing. If instead we assume the reductions in DI claiming are not simply a re-timing, then the reduced government costs from these fiscal externalities would reduce the net cost of \$1 of UI benefits by \$0.33 for benefit increases and \$0.01 for duration increases.

1.03 for benefit increases and 0.45-0.83 for duration increases.

III.G Top Marginal Tax Rates

Finally, we turn to the MVPF of top marginal tax rate changes. This example illustrates how we can utilize estimates from existing literature that attempts to provide empirical guidance on optimal government policy (e.g. optimal top tax rates, optimal unemployment insurance benefits, etc.).[44]

There is a large theoretical and empirical literature discussing the optimal top marginal income tax rate, summarized in Saez et al. (2012). This literature notes that any tax cut that provides \$1 of additional after-tax income is valued at \$1 by the mechanical beneficiaries. As a result, measuring WTP is straightforward. The cost to the government of the tax policy is more difficult. The cost of a tax cut that provides \$1 of benefits in the absence of a behavioral response is given by $1 + FE$, where FE is the impact of the behavioral response to the tax cut on government revenue.

For top marginal tax rate reductions, Saez et al. (2012) and Diamond and Saez (2011) show that this FE can be expressed as $\frac{\tau}{1-\tau}\alpha\epsilon^{ETI}$, where α is the Pareto parameter of the income distribution[45] and $\epsilon^{ETI} = \frac{1-\tau}{E[y]}\frac{dE[y]}{d(1-\tau)}$ is the elasticity of taxable income for top earners with respect to the top marginal "keep" rate of $1 - \tau$.[46]

The elasticity ϵ^{ETI} has been estimated using various tax reforms including the 1981 and 1986 tax decreases and 1993 increases in the top marginal income tax rate. We compute the MVPF of the historical tax policy changes that allowed researchers to identify ϵ^{ETI}. The MVPF for each tax reform is the ratio of WTP to cost, $1/(1 + FE)$:

$$MVPF = \frac{1}{1 - \frac{\tau}{1-\tau}\alpha\epsilon^{ETI}} \tag{8}$$

We translate estimates of ϵ^{ETI} estimated from five major tax reforms in 1981, 1986, 1993, 2001, and 2013, each of which is outlined in Appendix F.

To take one example, consider the 1981 tax cut that reduced the top marginal income tax rate from 70% to 50%. Saez (2003) finds an estimate of $\epsilon = 0.311$. We estimate $\alpha = 2.299$ from Atkinson et al. (2011). We plug these into equation (8). We use marginal tax rates of $\tau = 75\%$ and $\tau = 55\%$ before and after the reform, which include a 5% state tax adjustment. Combining, and averaging FE obtained using the pre-reform and post-reform tax rates, we obtain $FE = \frac{\tau}{1-\tau}\alpha\epsilon^{ETI} = 1.51$.

[44] While those literatures often consider the policies in isolation (e.g. optimal UI policy), we can translate the estimates into their implied MVPF to facilitate comparisons across policy domains.

[45] Mathematically, $\alpha = \frac{E[y_i]}{E[y_i - \bar{y}|y_i \geq \bar{y}]}$ where $\bar{y}$ is the threshold over which the top marginal income tax rate applies.

[46] Appendix F provides a derivation.

This means that the 70% marginal tax rate appears to have been on the "wrong side of the Laffer curve" so reducing tax rates may have increased revenue. In other words, the MVPF is infinite and the tax cut "pays for itself." However, it is important to note the statistical uncertainty in this estimate: we cannot reject an MVPF of 1 or ∞.

In contrast, for later reforms we find lower MVPFs. For example for the 1993 tax increase we find an MVPF of 1.85 (95% CI of [1.21, 4.03]). This distinction is not because of differences in ϵ, but rather it results from the fact that τ was much lower in 1993 than it was in 1981.

Comparison to "optimal" top tax rate. To compare our results to the literature on the "optimal" top tax rates, it is helpful to consider the case studied in Diamond and Saez (2011) where society is assumed to place no weight on the additional consumption of the rich. If the social welfare weights, η_i, are equal to zero for top earners, then the optimal tax is set to maximize government revenue: τ is chosen to be at the peak of the Laffer curve. This occurs when taxes are set so that the net cost to the government of providing a tax cut is zero, or $FE = -1$.

This approach then makes the additional assumption that the elasticity, ϵ^{ETI}, and α do not change when the tax rate changes. Solving for the optimal tax rate then implies $\tau^* = \frac{1}{1+\alpha\epsilon^{ETI}}$.

For $\alpha = 2.299$ and $\epsilon^{ETI} = 0.311$, this implies $\tau^* = 58\%$ inclusive of state and federal tax rates. The fact that this number is slightly below 70% is consistent with our finding of an infinite MVPF for the 1981 reform, in which tax rates were around 70%. In contrast to this optimal tax approach, the MVPF does not impose an assumption that society places no weight on the consumption of the rich.

IV Main Results: Targeting Kids versus Adults

We construct the MVPF for each of the policies in our sample. Here, we present all our baseline MVPF estimates and outline our main results. As noted, details on our MVPF constructions are provided in Appendices A-F.

IV.A Kids

We begin our discussion with the MVPFs of policies targeting children. Figure III presents the MVPF for each policy on the vertical axis plotted by the average age of the beneficiaries of the policy on the horizontal axis.[47] Each dot represents the MVPF of a particular policy, with labels

[47] In cases where both parents and children are beneficiaries of the policy, we assign the age of the "economic" beneficiary based on who has the highest WTP. For example, when analyzing the MTO experiment, which provided

provided in Table 1.

The figure reveals our primary result: direct investments in children have historically had the highest MVPFs, often paying for themselves. In addition to the evidence on the Medicaid expansions and admission to Florida International University, both discussed above, we also find high MVPFs for other education and child health policies. For example, Wherry et al. (2018) document that the discontinuous Medicaid coverage eligibility for children born after September 30, 1983 led to reduced medical costs and chronic conditions in adulthood. In Appendix D, we calculate that the upfront costs are fully repaid in the long run from reduced Medicaid and uncompensated care costs, leading to an infinite MVPF. More generally, all four major health insurance expansions to children studied in the past 50 years have MVPFs in excess of 10, with three of them paying for themselves.[48]

In addition to health policies, we find large MVPFs for education policies. The widely studied Perry Preschool program has an MVPF of 43.61; the more expensive Abecedarian model has an MVPF of 11.89 (neither of these estimates are statistically distinguishable from ∞).[49] In contrast to the idea that the returns to human capital investment diminish rapidly with age (Heckman, 2006), we find there is potential for high MVPFs investments throughout childhood. We find an infinite MVPF for increased K-12 spending due to school finance equalization as studied in Jackson et al. (2016).[50] We also find infinite MVPFs for several college policies, such as admissions to FIU and the provision of CalGrants to low income students. A key insight of our results is that many policies targeting children do not face the classic budgetary tradeoff. Instead, those expenditures pay for themselves in the long run.

Before drawing too many conclusions about each data point in Figure III, it is important to also

housing vouchers and counseling to parents with children, the age shown is the average age of the children in the household. This is because the policy induced higher earnings amongst the children, leading them to have a higher WTP for the policy than their parents.

[48]The only policy that does not have an infinite MVPF is the introduction of Medicaid. For this policy, we directly incorporate MVPF estimates from Goodman-Bacon (2017). His sample includes estimated impacts through age 55; our back-of-the-envelope calculations suggest that it is likely that forecasting these effects through 65 would lead the policy to pay for itself as well.

[49]To harmonize these estimates with other programs, we do not include the benefits to the government from reduced crime. This is both because these costs are difficult to quantify and most papers do not estimate impacts on crime outcomes. If we include a forecast of reduced government spending on the criminal justice system and policing, our point estimates suggest Perry Preschool paid for itself. However, the standard errors of these estimates also significantly increase. Including these costs for Abecedarian also increases its MVPF, but the policy does not appear to pay for itself.

[50]It is important to note that we only analyze one paper on K-12 education spending because of limitations in existing evidence on long-term outcomes. While there is a large literature looking at the impact of schools spending on test scores, we lack a reliable method to translate these effects into long-run impacts. Jackson et al. (2016) demonstrate the potential for high returns to K-12 education, but future work is needed to robustly establish the presence of high returns to K-12 investment.

note there is sampling uncertainty inherent in our estimates. Figure IVA plots each MVPF along with its 95% confidence interval. In some cases, our estimates are relatively precise. For example, both the Medicaid expansion to pregnant women and infants and admissions to FIU have confidence intervals that reject any finite MVPF. We can rule out any positive net cost to the government.[51] In many other instances, however, the conclusions at the individual policy level are less clear due to the sampling variation in the underlying estimates. For example, the 1990 healthcare expansion to children born after Oct 1, 1983 has a confidence interval ranging from 0.26 to infinity. In other words, we cannot with 95% confidence reject the hypothesis that the policy paid for itself, nor can we reject the hypothesis that the policy provides much less than $1 of benefits per dollar of government spending.

To reach more precise conclusions at a broader level, we pool across policies using category averages. We imagine a new policy that spends $1 of initial program cost on each policy j in category J containing N_J policies. We then construct the MVPF of this category-average policy as:

$$\overline{MVPF}_J = \frac{\frac{1}{N_J}\sum_{j\in J}\frac{WTP^j}{C_j}}{\frac{1}{N_J}\sum_{j\in J}\left(1+\frac{FE_j}{C_j}\right)} \tag{9}$$

where the numerator is the average willingness to pay per dollar of program cost and the denominator is the average net cost to the government of the category-average policy.[52]

Figure IVB presents the category-average MVPFs. On average, spending on child education, child health insurance, and college policies have historically had high or infinite MVPFs. One dollar of spending across each of the policies in each of these categories has an infinite MVPF in child education (95% CI of $[12.1, \infty]$), ∞ in child health (95% CI of $[25.0, \infty]$), and ∞ in college policies (95% CI of $[4.5, \infty]$).

We can dig deeper into these category averages by focusing on the net costs to the government of these policies (the denominator in our formula in equation (9)). Figure V computes the average net cost to the government per $1 of programmatic expenditure spent evenly across the policies in each category. This allows us to explore the extent to which different types of policies have paid for themselves. For example, $1 invested in the four major Medicaid expansions to children has paid back an estimated $1.78. In other words, the spending actually generated $0.78 of surplus to

[51] This analysis takes the reported point estimates and standard errors as given. In Section IV.D, we discuss how these conclusions are affected by publication bias. We note that these estimates appear to suffer from significant bias, our broad conclusions hold true after using the correction methods of Andrews and Kasy (Forthcoming) .

[52] We construct this average measure, as opposed to a precision-weighted average or other measure, because it corresponds to a feasible policy at the time of initial implementation. It is straight forward for the government to construct a policy that spends an equal amount on each of these programs.

the government in the long run.[53]

Having established this primary result, it is important to qualify that these patterns do not hold *uniformly* across policies. There is considerable variation in MVPFs from policy to policy. For example, we find lower MVPFs ranging from -0.23 to 1.48 for job training policies, such as an estimate of 0.15 for Job Corps – a program targeted towards at-risk youth.[54] We also analyze 14 examples of college policies where the MVPFs fall below 2.[55] In most cases, this is because those policies represent transfers to existing students, rather than expenditures that increase attainment.[56] In some cases, expenditures may even negatively impact student attainment. For example, Cohodes and Goodman (2014) analyze the impact of the Adams Scholarship in Massachusetts. They find that this merit aid program does not induce more students to go to college. Rather, it induces individuals to change colleges to attend schools where they are eligible to use the Adams scholarship. The change in schooling actually results in a fall in graduation rates.[57] Incorporating these schooling declines, we calculate that the program has an MVPF of 0.72. Job training or education polices like this one do not substantially increase human capital and so they do not recoup meaningful portions of their initial costs via higher tax revenue.

We also find lower MVPFs for transfers to disabled children, such as an MVPF of 0.76 for expanded eligibility for Supplemental Security Income (SSI) at age 18 analyzed in Deshpande (2016). It is important to note that spending on these policies may increase social welfare, even though they have lower MVPFs. Decisions about optimal policy are determined by the welfare weights that the government places on policy beneficiaries. If the government makes it a priority to provide support for disabled children, these SSI expansions may be welfare-enhancing.

[53]Analogously, Appendix Figure II presents willingness to pay per dollar of programmatic spending. For our baseline WTP measures, we find very similar patterns: much higher estimates of $\frac{1}{N_J}\sum_{j\in J}\frac{WTP^j}{C_j}$ for child policies than for policies targeting adults.

[54]The one potential exception to this is the recent Year Up RCT, analyzed in Fein and Hamadyk (2018) who document large increases in earnings in the two years after initial implementation. As we discuss in Appendix C, if these earnings gains persist for an additional 5 years the MVPF would be 2.78, and if they persist for 21 years the MVPF would be infinite.

[55]Our analysis also demonstrates the limitations of the traditional way that research papers report the impact of college expenditures. It is very common for papers to note the percentage point increase in enrollment associated with a $1,000 in expenditures. The difficulty with that approach is that it doesn't account for the number of inframarginal students receiving the benefit. Providing $1,000 to 10% of the school-age population to achieve a 3.6 percentage point increase in enrollment may be a very efficient investment, while providing $1,000 to 80% of the school-age population to achieve a 3.6 percentage point increase is mostly a transfer to existing students. For this reason, there are cases where we find substantially different MVPFs for policies that had similar percentage point enrollment effects.

[56]In Section IV.C below we discuss how our results on college expenditures vary with the method of our MVPF calculation. While we find persistently high MVPFs when long-run earnings outcomes are observed, we find lower MVPFs when we project earnings gains from attainment outcomes.

[57]The authors argue that, on average, the Adams-eligible schools provide a lower quality education.

IV.B Adults

In contrast to policies targeting children, we generally find lower MVPFs (e.g. 0.5-2) for policies targeting adults. For example, in contrast to the near infinite MVPFs for child health insurance expenditures, we find MVPFs ranging from 0.40 to 1.63 for the six health insurance policies in our baseline sample targeted to adults.[58] Along the same lines, we find MVPFs ranging from 0.43-1.03 for unemployment insurance policies, 0.74-0.96 for disability insurance expansions, and 1.12-1.19 for earned income tax credits. We find MVPFs of housing vouchers ranging from 0.76 using an RCT of the provision of housing vouchers to families on cash welfare (Mills et al., 2006) to 0.65 using assignment of vouchers in Chicago via lottery (Jacob and Ludwig, 2012).

The lower MVPFs reflect the fact that many of these expenditures have been shown to reduce labor earnings through labor market distortions. As depicted in Figure V, the average cost per $1 of government spending on these adult policies is generally slightly above $1. This result contrasts with our findings on expenditures direct toward children, for whom labor market earnings tended to rise, leading to a decline in net costs. There are a limited number of cases, such as the Job Training Partnership Act and National Supported Work Experiment, where investment in adults sought to increase earnings by increasing human capital. Those policies, however, did not produce persistent earnings gains, and so they still yield relatively low MVPFs. Their values range from 0.15 to 1.48.

As with our main results for policies targeting children, these findings represent general patterns. They do not hold uniformly across all policies targeting adults. In particular, there are two types of adult policies that tend to result in higher MVPFs: reductions of high marginal tax rates for top incomes and policies with indirect spillovers onto children.

Top Tax Rates We find high MVPF point estimates for historical reductions in the top marginal tax rate when the initial tax rate lay at 50% or above. In the case of the 1981 reform, the tax bill reduced the top federal marginal tax rate on income from 70% to 50%. Using estimates of the elasticity of taxable income from Saez 2003, we calculate that the MVPF is ∞ (95% CI of $[0.94, \infty]$). This implies that marginal tax rates were beyond the top of the "Laffer curve" prior to 1981. Our confidence interval, however, suggests this estimate contains considerable sampling

[58] Those adult health insurance estimates include expenditures such as the subsidies in the Massachusetts health insurance exchange prior to the Affordable Care Act. In that case, Finkelstein et al. (2019) exploit discontinuities in the subsidy schedule to estimate both individuals' willingness to pay for insurance and the cost those individuals impose on the government. Translating these estimates into an MVPF suggests values ranging from 0.800 to 1.09 for different subsidy eligibility levels.

uncertainty.[59] Along the same lines, we analyzed the 1986 reform and found an MVPF of 44.3, with a confidence interval ranging from 2.37 to ∞.

While this may be considered by some to be suggestive evidence for Laffer effects in tax policy, it is important to approach that conclusion with considerable caution. In the case of the 1981 reform our confidence intervals suggest we cannot rule out an MVPF close to 1. In other words, we cannot rule out the conclusion that the policy produced no positive fiscal externality.[60] As compared to these findings on taxes, our results suggest stronger evidence for the presence of Laffer effects when investing in young children.

Spillovers onto Children We also find that spending on adults may have high MVPFs if those policies have spillover effects on children. For example, Chetty et al. (2016) study the long-run impact of the MTO experiment, which gave families residing in public housing projects a voucher and counseling to assist them in moving to lower poverty neighborhoods.[61] Chetty et al. (2016) document that the program significantly increased later-life earnings for young children, but they find null or even slightly negative impacts on earnings for children who were teenagers at the time their parents obtained the vouchers. Combining these effects across all subgroups suggest the impacts on the young children outweigh the adverse impacts on the older children, leading to an infinite MVPF.[62] This high MVPF is driven solely by child outcomes, as the policy has no significant effect on economic outcomes for adult beneficiaries.

One policy with a substantial degree of uncertainty about potential spillovers onto children is the EITC. Appendix Figure IIIC shows how our MVPF estimates would change if one attempted to impute effects on children using different estimates from previous literature. In particular, we take the MVPF for the 1993 OBRA tax reform and supplement that estimate with spillover effects of the

[59] As we discuss in Appendix F, these estimates appear to have considerable uncertainty not just from sampling uncertainty but also model uncertainty: using different taxable income estimates from existing literature studying these reforms can generate wide variation in the MVPFs of these tax reforms, preventing precise conclusions about their MVPFs.

[60] It is also worth noting that more recent reforms have substantially lower MVPFs. The MVPF of the 2013 top tax rate changes is 1.16.

[61] Because the program was targeted to families already in public housing and because the cost of public housing is similar to the cost of a voucher, the primary marginal cost of the program was the cost of the counseling (roughly \$3,783 per family).

[62] Not all policies providing benefits to parents generate such large spillover effects onto children. For example, Price and Song (2018) find that the Negative Income Tax experiment led to a reduction in children's earnings in adulthood, which partially explains its low MVPF of 0.11. In other cases, such as the provision of housing vouchers in Chicago, and the provision of housing vouchers to families on AFDC and the expansion of AFDC benefits, there is suggestive evidence that positive spillovers on children are small. In those instances, researchers have documented that the policies have limited effects on outcomes such test scores, college attendance, and birthweight. Appendix Figure IIIA presents results for policies in our baseline sample where child impacts are observed. Panels B and C show how the MVPFs change when impacts on children are incorporated or removed from the MVPF calculation.

EITC estimated in other contexts. Projecting earnings impacts based on child test scores produces MVPFs that range from 3.48 to ∞, while incorporating effects on college attendance produces MVPFs from 1.15 to 1.73. (The college effects are restricted to a small subset of recipients, so it is unsurprising that the MVPFs remain small.) Incorporating the work of Bastian and Michelmore (2018) on long-term earnings would result in an infinite MVPF, suggesting the policy pays for itself.[63]

This uncertainty highlights the importance of understanding the potential spillovers onto children. It also reinforces our conclusion that policies which raise children's human capital often have the highest MVPFs. We return to this issue in Section VI.A, where we use the MVPF framework to quantify the value to governments of more precise estimates for potential long-run effects of policies on children.

IV.C Robustness

The creation of these MVPF estimates inevitably requires that we make a number of judgment calls regarding both the set of causal effects included and the methodology used to translate those effects into an MVPF. Here, we illustrate the robustness of our main conclusions to these assumptions.

Discount Rates and Tax Rates Constructing the MVPF for policies with dynamic effects requires a choice of a discount rate. Appendix Figure IV illustrates our results are robust to a range of different discount rates between 1% and 10%, which span the real interest rates observed in recent history. For illustration, Figure VI provides an estimate of the MVPFs for an interest rate of 7% instead of our baseline 3% assumption.[64] The clearest discernible change is a drop in the MVPF for childhood education from an infinite MVPF to an MVPF of 4.05. In contrast, college education policies continue to have high MVPFs. Intuitively, when investments take a long time to pay off, they have MVPFs that are more sensitive to interest rates.

In addition to discount rates, Appendix Figure V shows our results are robust to alternative assumptions about the tax rate faced by individuals. For example, assuming only a 10% tax on

[63] We exclude these results from the baseline estimates because Bastian and Michelmore (2018) do not estimate the effect of a particular EITC expansion, but rather pool across many state and federal policy changes. In Appendix F, we note the impact of incorporating their estimates. The fact that these impacts matter is consistent with our broader conclusions that potential spillovers onto children can generate high MVPFs for adult-targeted policies.

[64] We use a baseline 3% discount rate in analyzing all of our policies, but one could also consider using current government borrowing rates to conduct these calculations. The intuition there is that many government expenditures are deficit financed rather than part of budget neutral policy packages. In that context, the long-term net cost of government expenditures are determined in part by the governments cost of borrowing. If one took that approach to determining the discount rate, declines in the real interest rate over the last 40 years would imply that set of high return investments has risen over time.

earnings continues to suggest high MVPFs for child health and education policies. In the other direction, Appendix Figure VD shows that if we raise the tax rate assumption to 30%, our primary conclusions are even stronger. The increased earnings induced by the childhood investments result in even larger increases in government revenue. This higher tax rate is not implausible because our CBO-based tax and transfer rate excludes payroll taxes, housing vouchers contributions, and cash welfare programs.

Fixed Forecast In addition to a 3% interest rate, our baseline forecast method assumes counterfactual incomes remain proportional to the population mean income throughout the life cycle. A more conservative approach is to assume the counterfactual income profile does not grow at all over the life cycle, but rather the impact of the policy on earnings remains fixed in all projected years. Figure VIC repeats Figure IVB for this alternative projection procedure and finds similar results.

Willingness to Pay Method Figure VID presents the MVPFs for an alternative "conservative" WTP specification for each policy.[65] Here again, we find large MVPFs for policies targeting children. This is largely a mechanical result of the pattern illustrated in Figure V, where we report the net costs of each policy. Even if policies have a low WTP, they may still have a very high MVPF if the government recoups much of its initial expenditure.[66]

Dropping College Forecasts Our baseline sample includes some policies targeting children for which impacts on income are not directly measured. Most notably, we include college policies where researchers have observed a measure of attainment such as initial enrollment, college credits or degree receipt. In those cases we forecast income impacts using estimates from Zimmerman (2014) on the returns to college. (Details of this method can be found in Appendix I.) In Figure VIIA we present the results for our "restricted" sample that drops all policies where causal effects on income were not measured. We find meaningfully higher MVPFs in this restricted sample, with policies repaying on average $0.56 for each dollar of initial investment. By contrast, our net cost calculations in the baseline sample suggest college policies repay $1.38 for every dollar invested.

[65] The examples in Sections III.A and III.B provide details on this definition for the admissions to FIU and Medicaid expansions to pregnant women and infants. More details on this approach for all other programs can be found in the Appendices A-D explaining each of our MVPF calculations.

[66] In cases where policies increase human capital but do not have an infinite MVPF, this apporach substantially reduces the estimated bang for the buck. As we detail in Appendix B, most college expenditures with positive net costs see substantial MVPF declines using this approach. The infinite MVPF in the group average is driven by the presence of high return policies where earnings effects are observed.

Exploring the restricted sample illustrates the difficulty of explaining this pattern. It is no surprise that the MVPF of admission to FIU is quite large. The policy was directly focused on increased admissions for academically marginal students, and so nearly all beneficiaries were newly induced into four-year college enrollment. (Some attended FIU instead of community college while others attended FIU instead of receiving no college education.) This means the expenditures were directed toward increasing human capital, rather than merely providing a transfer to students who would have enrolled anyway. That said, the other two college policies in the restricted sample have high MVPFs despite more modest increases in attainment (Bettinger et al., 2019; Denning et al., 2019).

The higher MVPFs in cases where earnings are observed may be due to several factors. It could be that follow-up earnings estimates were obtained in cases where researchers believed that outsized earnings gains were present. In that case, our baseline sample would be a more accurate representation of the average returns to college spending. By contrast, it could also be that college expenditures increase earnings through channels other than attainment.[67] In that case, our attainment-based projection method could systematically understate the MVPF of these policies. It is clear that more work is needed to investigate the relationship between projected and observed outcomes.

Long-Run Estimates In many cases, our MVPFs for policies targeting adults rely upon estimates of short-run earnings impacts. Consequently, one might be worried that our low MVPFs for adult policies are driven by policies for which we do not observe long-run impacts. For example, perhaps a subsidy to adults distorts labor earnings in the short run, but has positive impacts in the long run that are not always measured. To assess this, we restrict attention to the subset of policies for which we observe at least five years of income estimates. In general, we continue to find MVPFs in the 0.5-2 range for these policies, as we illustrate in Figure VIIB. For example, we find MVPFs of 0.74 for the disability insurance (whom French and Song, 2014 observe for at least 10 years after random assignment to a judge[68]) and 1.48 for the National Supported Work experiment that provided job placement and training to women on Aid to Families with dependent children (whom Couch (1992) observe for 9 years after random assignment).

[67] Bettinger et al. (2019) argue that other channels such as graduate school attendance may help explain these earnings gains, while Denning et al. (2019) argue that channels other than degree completion are unlikely to explain their results.

[68] The primary results in French and Song (2014) report earnings three years over three years after treatment, but graphical evidence shows the earnings declines remains the same or dissipates slightly over 10 years.

Quality of Estimates / Validity of Identification Assumption One concern with creating MVPFs based on existing literature is that the causal effects underlying the MVPFs may be of varying quality. In order to address this, we draw upon prominent survey articles to determine the sample of policies we analyze. Our hope is that this approach helps to provide a screening mechanism to ensure that the underlying causal effects are reasonable. Here, we provide additional evidence that the quality of underlying estimates do not influence our results.

In Figure VIIC, we restrict our sample to MVPF estimates where the primary causal estimates come from a randomized controlled trial, lottery, or a regression discontinuity design. While this significantly limits the sample, the broad pattern of results remain the same. In Figure VIID, we perform a similar test, restricting our sample to peer-reviewed publications. We find that our confidence intervals expand, but our basic results remain the same. Direct childhood investments continue to have the highest MVPFs.

IV.D Publication Bias

All of the robustness analyses above take the estimates from existing literature as given. However, one might also be concerned that the research and publication process suffers from the problem of publication bias, where studies are published only if they find clear positive (or negative) effects. In particular, one might worry that research on children is more likely to be published if it finds statistically significant positive effects on children in adulthood. Conversely, one could imagine that research on adults is more likely to be published if it finds statistically significant evidence of distortionary or negative effects on adult outcomes. Even if the programs studied had no true effect, all estimates have sampling uncertainty. Selective publication of those uncertain estimates could produce a biased picture of the true effect, especially in cases where we rely upon studies with small samples.

To address this, we use the approach in Andrews and Kasy (Forthcoming). This paper provides a method to both test for publication bias and to correct for the impact of publication bias on the observed set of estimates. Andrews and Kasy (Forthcoming) note that if one knows the relative publication probabilities, one can adjust the estimates to remove the effects of publication bias.

Testing for Publication Bias We first test for the presence of publication bias in our estimates using the full set of causal estimates employed in our baseline MVPF construction, and we also explore heterogeneity in the presence of publication bias amongst sub-samples of the data.

Throughout, we test separately for publication bias for estimated impacts on children and parents. We test for bias at the negative and positive $p > 0.10$ and $p > 0.05$ thresholds.[69]

Table III presents the results. Columns (1) and (4) present results testing for bias at the .10 threshold for children and adults; Columns (2) and (5) test for bias at the 0.05 threshold; and Columns (3) and (6) test simultaneously for bias at the .10 and .05 thresholds. The numbers reported represent the likelihood of being published relative to if the study had no statistically significant effect.

The results suggest evidence of publication bias, but with substantial heterogeneity, especially for impacts on adult outcomes. In the baseline sample we find studies of child outcomes are 3.69 times more likely to be published if they find positive effects on children with $p > 0.10$ relative to a finding of no statistically significant effect.

In contrast, we find no evidence in the baseline sample that studies with adult outcomes are more likely to be published if they show significant results. However, this finding of a lack of publication bias is not robust to restricting our sample to effects that are estimated with less statistical precision. To see this, the remaining 6 columns of Table III repeat these estimates for a restricted sample of estimates with t-stats below 10. Here, we find that estimates showing negative or distortionary impacts for adults with $p < 0.10$ are over ten times more likely to be published relative to insignificant results.

To place these estimates in perspective, Andrews and Kasy (Forthcoming) use data from Camerer et al. (2016) to estimate that lab experiments in economics are 26-35 times more likely to be published if they have statistically significant results; in contrast, they estimate that studies of the minimum wage are 3 times more likely to be published if they are statistically significant. This means the publication bias we find for both adult and child outcomes in our sample falls between the publication bias estimated in these minimum wage and lab experiment studies.

Correcting for Publication Bias If one knows the likelihood of publication given the estimates, Andrews and Kasy (Forthcoming) show that one can adjust the estimates to remove the effects of publication bias. Loosely, if estimates with $p < .05$ are more likely to be published, then less precise estimates that are close to that 0.05 threshold will be adjusted towards zero to account for the fact that these may have arisen by chance and the estimates were disproportionately likely to be published as a result.

[69] For each policy, we input all estimated causal effects into the test; we exclude program parameters such as take-up rates and first stage coefficients in IV identification designs.

We adjust for publication bias in two ways. First, we use the estimated publication bias likelihoods in Table III and assume that all child and adult estimates are subjected to this publication bias in our baseline sample. For each causal effect, we translate the estimated effect and its sampling distribution into estimates and sampling distribtuions that are unbiased from the distorted likelihoods of publication using the methods of Andrews and Kasy (Forthcoming). We then use the resulting estimates of the true and biased sampling distribution to form the MLE estimate of the true effect given our observed estimate.[70] This approach provides a bias-corrected MLE estimate of the MVPF.

Figure VIIIA presents the results using the bias correction implied by the bias threshold model reported in columns 3 and 6 of Table III for child and adult estimates. Broadly, the results show that correcting for this estimated degree of publication bias does not affect our primary conclusions. We find a slight decrease in the MVPFs for child education policies, such as preschool programs. We continue to find large finite or infinite MVPFs on child health and college policies.

While this first approach assumes that all estimates are subjected to the estimated amount of publication bias, it could be the case that some types of estimates are subject to more publication bias than others. Indeed, as documented in Table III, we do find evidence of heterogeneity in the degree of publication bias in our sample. Unfortunately, a full exploration of subgroup heterogeneity is limited by the small samples of our data. For example, we only observe two preschool programs in our baseline sample (Abecedarian and Perry Preschool), and so we cannot determine if preschool estimates are subject to a unique degree of publication bias.

In order to account for the potential of greater publication bias, we proceed by imagining that our estimates had been subjected to a larger degree of publication bias than we observe in the actual sample. In particular, we correct for publication bias under an assumption that our observed estimates for the impacts of child policies suffered as much publication bias as has been documented in small-sample experimental economics studies. To that aim, Figure VIIIB presents corrected estimates under an assumption that statistically significant estimates of positive impacts on children are 35 times more likely to be published. Figure VIIIB show that under this specification, our primary conclusions still hold – policies targeting children have the highest, often infinite MVPFs. This means that even if the literature on children is subjected to much larger

[70] We thank Isaiah Andrews and Max Kasy for their invaluable guidance, and in particular for suggesting this approach to construct an estimate of the bias-corrected MVPF. The MLE approach remains valid under nonlinear transformations. One limitation we acknowledge is that we treat the MLE estimates of each estimate as independent, and do not account for correlation across estimates within the same study.

degrees of publication bias than seems to be implied by the patterns in the data, there is still clear evidence of high MVPFs for investments in low-income children.

IV.E Summary

Broadly, we find that direct investments in low-income children have had the highest, often infinite, MVPFs. The MVPFs are high throughout childhood, rather than declining with age.[71] In contrast, we find lower MVPFs around 0.5-2 for policies targeting adults. These findings represent general patterns, but there exist exceptions where child policies have low MVPFs (e.g. SSI and job training) and adult policies have high MVPFs (e.g. top tax rates changes and in-kind transfers with spillovers). The results are robust to a range of alternative assumptions and sample criteria, and they are robust to adjusting for the presence of publication bias.

V Mapping the MVPFs to Theory

The MVPF provides an empirical method for evaluating the effectiveness of different policies for improving social welfare. Having established the key patterns of the data, it is natural to place our empirical results into the context of theoretical literature on optimal government policy. In this section, we outline how our results speak to that theory.

V.A Redistribution Amongst Adults

The MVPF measures the price of redistributing to different policy beneficiaries. In this sense, it is most closely related to the theoretical literature on optimal taxation and redistribution. The seminal work of Mirrlees (1971) and the large subsequent body of theoretical tax literature analyzes the optimal progressivity of the tax schedule and the optimal method of redistribution (e.g. a negative income tax versus an earned income tax credit). In general, the theory suggests that a progressive planner should be willing to incur efficiency losses to move resources from the affluent towards the lower regions of the income distribution. The MVPF provides an empirical means of testing that basic prediction.

As previously explained using the Okun's bucket logic, the ratio of MVPFs across two different tax changes measures the price of moving money between the respective beneficiaries. This means that, in the case of a progressive planner, the MVPF of tax changes should increase with income.

[71] It is important to be clear that while our estimates suggest high returns to policies investing in older youth, the policies in our sample affect a range of sub-populations. As a result, further work is needed to assess how the rate of return on investment varies for a given child over the life cycle.

Figure IXA shows the MVPF of each tax policy change we analyze and the income levels of the associated beneficiaries. Consistent with this prediction, we observe a loosely upward slope.

For example, the 1993 tax reform (OBRA93) simultaneously raised top marginal tax rates and expanded EITC. The MVPF of the increased top tax rates led to an MVPF of 1.85 (95% CI of [1.21, 4.03]), and the expansion of EITC led to an MVPF of 1.12 (95% CI of [0.27, 1.20]). This suggests the tax schedule created under the 1993 reform is optimal if one is indifferent to providing $1.85 to top earners versus $1.12 to those on EITC. To the extent one's social preferences strictly prefer $1.12 to low-earners (or strictly prefer $1.85 to top earners), our results suggest more progressive (regressive) taxation than the 1993 schedule would be optimal.[72]

In addition to broad redistribution from the top to bottom of the income distribution, the MVPF also provides insights into the optimal method of redistribution amongst low-income beneficiaries. We again see potential evidence of an upward slope within low-income beneficiaries, as shown in Figure XIA. For example, the 1.12 for the 1993 EITC expansion and 1.20 for the 1986 EITC expansion are above the 0.11 MVPF point estimate for the Negative Income Tax experiment (NIT). While the EITC has historically been an efficient method of redistribution, it targets more affluent beneficiaries than the NIT. One may still prefer the NIT if one prefers $0.11 to those at the very bottom of the income distribution relative to just over $1 to those with earnings in the EITC range.

V.B Testing Atkinson-Stiglitz: When Are Transfers In-kind More Efficient than Cash?

Our results quantify the cost of additional redistribution through the tax schedule. The low MVPFs for policies targeting very low-income households such as NIT raises the question of whether other methods of redistribution - perhaps through in-kind transfers - can be a more effective than cash.[73] There is a large theoretical debate on this question, which largely centers around the applicability of the Atkinson-Stiglitz theorem (Atkinson and Stiglitz, 1976; Hylland and Zeckhauser, 1981). When utility satisfies a "weak separability" assumption, redistribution in cash is always more desirable

[72] Our estimate for the MVPF of the 1993 EITC is based on evidence from Meyer and Rosenbaum (2001) on the fiscal externality associated with the labor supply responses of single women. It is worth noting, however, that there is considerable debate over the fiscal externalities associated with the EITC. On the one hand, several recent papers have argued that reductions in transfers have offset a substantial portion of the cost of historical EITC expansions (Hoynes and Patel (2018);Bastian and Jones (2018)). These large fiscal externalities can produce an MVPF in excess of 4 (Bastian and Jones, 2018). On the other hand, recent debates have argued the effects are overstated in the current literature because the impact of the EITC expansions cannot disentangled from the effects of contemporaneous welfare reforms (Link – Twitter Discussion). These conflicting estimates suggest there is a high value to future work that reconciles these findings.

[73] See Currie and Gahvari (2008) for a summary of the theoretical and empirical literature debating the desirability of these transfers relative to cash transfers.

than in-kind. However, when the weak separability assumption is violated, such as cases where an in-kind transfer has a less distortionary impact on earnings than a cash transfer, then the in-kind transfer may be preferable to a distributionally equivalent tax cut.

We can test the empirical applicability of Atkinson-Stiglitz by comparing the MVPFs of tax policy changes to the MVPFs of in-kind transfers. In the case of in-kind transfers, we focus on housing and food policies, two of the largest sources of such transfers in the US. Figure IXB shows the MVPF estimates for both in-kind transfers and cash transfers/tax credits, reported as a function of the beneficiaries' incomes. The Atkinson-Stiglitz theorem would suggest that the MVPF for an in-kind transfer would fall below the MVPF of a cash transfer or tax credit targeted to beneficiaries at the same income level.

We find some evidence consistent with the predictions of the Atkinson-Stiglitz theorem, although the estimates are statistically imprecise. Both the housing vouchers in Chicago (Jacob and Ludwig, 2012; Jacob et al., 2014) and the provision of Welfare to Work housing vouchers (Mills et al., 2006) have MVPFs below that of distributionally equivalent tax cuts. That is because the vouchers have negative effects on short-term labor supply and no discernible impact on outcomes for the children of voucher recipients. By contrast, there are some prominent cases where the MVPFs of in-kind transfers exceed that of tax cuts. For example, the Movement to Opportunity (MTO) Experiment explained previously increased earnings of young children by a sufficient amount to pay for the cost of the in-kind policy (the policy had an infinite MVPF with 95% confidence interval of $[-2.80, \infty]$). Similarly, the spillover effects onto children for the introduction of food stamps policy leads to an MVPF of 1.04. Both point estimates suggest these policies are more efficient than cash transfers. However, in contrast to the traditional ways in which these policies are modeled, the violation of the weak separability assumption comes not from a short-term change in earnings, but rather the long-run indirect impact on children.

V.C Optimal Targeting

There is a large literature in optimal policy design focused on improving efficiency by targeting the right subset of individuals. In general, this work focuses on the use of "tags" – characteristics of program eligibility that are generally not manipulable (Akerlof, 1978).[74] As expressed in Weinzierl (2011), "The fundamental challenge for tax policy design, as first posed rigorously by Mirrlees

[74] If the tag were manipulable, then individuals not intended as beneficiaries of the policy could distort their behavior to obtain the benefit. To first order, they would not value the transfer by the envelope theorem, consequently lowering the MVPF of the policy.

(1971), is to use observable information to redistribute income to the low-skilled and pay for public goods without discouraging work effort." To that aim, previous literature has identified recipient age as a potentially valuable tag for optimal government policy.[75] To the extent to which work incentives evolve over the life cycle, policies that are age-dependent can increase welfare. Our results are consistent with that hypothesis, with perhaps one nuance being that the ages that matters are the ages of children in the household. Here, we discuss the implications of a couple prominent cases where existing literature has documented heterogenous impacts that vary based on age of children in the household: the Moving to Opportunity (MTO) experiment and the introduction of food stamps.

To begin, our analysis of MTO finds an infinite MVPF with a confidence intervals of $[-2.80, \infty]$. These numbers mask substantial variation in outcomes based on the age of the impacted children. Young children see far greater benefits from MTO. In families with children younger than 12, the MVPF is infinite with a confidence interval contained at infinity. In families with children older than 12, Chetty et al. (2016) find slightly negative effects argued to reflect disruption spillovers onto the children from moving, which leads to an MVPF for this subset of families of -3.55 with a 95% confidence interval of [-4.54, ∞].

In a similar vein, work on the introduction of food stamps by Bailey et al. (2019) documents the largest impacts on young children (e.g. ages 0-5). Our baseline MVPF of 1.04 discussed in Section III.C accounts for the fact that food stamps are provided to families with children of all ages. We can also consider a policy that targets these benefits to families with children aged 0-5. Assuming the behavioral responses of this targeted policy are similar, it would suggest this targeted policy would have an MVPF of 2.28.

In each of these cases, our point estimates suggest that the MVPF of these policies could be substantially improved with targeting based on the ages of children in the household. These patterns also suggest a value to future work identifying additional tags for which policies can have high spillovers onto children. More generally, they suggest a value of future theoretical work focused on not just the age of the beneficiary but the age of children in the household when considering optimal tagging policies.

[75] For example, see the large literature on the "New Dynamic Public Finance" that establish large gains from age-dependent taxation (e.g. Golosov and Tsyvinski (2007); Stantcheva (2017).

V.D Missing Markets: The MVPF of Social Insurance Policies

There is a large literature suggesting the government may have a role in correcting private market failures, such as those resulting from asymmetric information in insurance markets (Akerlof, 1970). Asymmetric information can lead to adverse selection (or, in the extreme, market unraveling) that prevents individuals from buying insurance in private markets. As highlighted by Akerlof (1970), these missing markets can occur even though willingness to pay for the insurance exceeds the marginal cost of its provision.

The MVPF provides an ideal way to assess the potential welfare implications of policies that correct market failures. It can even provide a test for these market imperfections. If private markets are inefficient and undersupply insurance, then the MVPF of government insurance expenditures should be in excess of one as willingness to pay exceeds net costs. In the absence of asymmetric information or other market failures, willingness to pay should not exceed cost because private markets should satisfy demand in all such cases.

We can take this intuition to the data to see if market imperfections provide a rationale for historical interventions in insurance markets. We find modest MVPFs for social insurance policies targeting adults with values often hovering around one. For example, we examine several expansions of adult health insurance and find MVPFs ranging from 0.80 to 1.63. This includes the 1960's expansion of Medicare Finkelstein and McKnight (2008), which lies at the upper bound of this interval with an MVPF of 1.63. It also includes the expansion of Medicaid eligibility in Oregon (Finkelstein et al.,Forthcoming), with a baseline MVPF of 1.16.[76] We see similar patterns in our evaluation of unemployment insurance with MVPFs ranging from 0.43 to 1.03. In these instances, we find some cases where the insurance value of the expenditure leads the willingness to pay to exceed net costs and others where it does not. Overall, the MVPFs are roughly close to one and we do not find that imperfections in these markets have led to extraordinarily high MVPFs.

The modest MVPFs for adult insurance policies lie in stark contrast to the MVPFs of insurance policies targeted toward children. As shown in Section IV, these children-focused policies have infinite or nearly infinite MVPFs. These large MVPFs are particularly interesting because the efficiency gains come from spillover increases in government revenue, rather than providing a product where willingness to pay exceeds the initial cost of provision. This suggests that, in the logic of

[76]These results vary from 0.50 to 1.16 based on the specification used. This variation illustrates the difficulty of comparing policies to a single benchmark such as an MVPF of one. The specification and sample uncertainty make it difficult to assess the relative efficiency of individual policies.

Akerlof, the most important "missing market" is not insurance for adults but instead a market for investment in human capital for children.

V.E Human Capital Policies

Our results are also directly related to the large theoretical literature studying optimal government policy in human capital markets. The literature often focuses on the optimal price of, or subsidy for, human capital investments. A primary insight is that because the government taxes incomes, it effectively taxes the returns to human capital. Consequently, subsidizing investment in human capital can help internalize these fiscal externalities and ensure that tax policy does not dis-incentivize investment. Under certain conditions, the optimal human capital policy is full deductibility of educational expenditures (Bovenberg and Jacobs, 2005; Stantcheva, 2017).[77] This ensures that private decisions about human capital investment are aligned with the socially optimal decisions that maximize government revenue.

Our findings are quite consistent with the intuitions provided by this literature. The primary reason we find large MVPFs for college policies is because of dynamic effects on human capital accumulation and future impacts on earnings. This suggests that government subsidies for human capital may be welfare-enhancing. That said, our results are not necessarily consistent with the precise suggestion that full deductibility of educational expenditures is the optimal price of human capital. There are a couple of ways in which our results present a more varied picture – one that perhaps questions whether optimal human capital policies should rely on a maximization program that assumes full information about the future returns.

First, we observe relatively low MVPFs for several policies meant to encourage college attainment via the tax code. We look at the tuition deduction (Hoxby and Bulman, 2016), the American Opportunity and HOPE Tax Credits (Bulman and Hoxby, 2015) and the combination of the HOPE and Lifetime Learners credits(Long, 2004). While these estimates are noisy, the point estimates do generally point to small effects at best. The contrast to the higher MVPFs of direct schooling investments suggests that households may not re-optimize their educational investments in response to these benefits.[78] In that case, subsidies to parents through the tax code may best be thought of as transfers to adults as opposed to policies that induce investment in human capital. This means it would be natural for these policies to have an MVPF near 1.

[77] Formally, the optimal price of \$1 of human capital is $\$1 - \tau$, where τ is the tax on labor earnings.

[78] As many have noted, these tax deductions and credits are granted after the original investment is made. Ex-post reimbursement may not be sufficient to generate optimal investment if credit or informational constraints are in place.

Second, in our extended sample we also examine several cases where "nudges" or "informational treatments" are used to encourage college enrollment (e.g. support filling out the FAFSA as studied in Bettinger et al., 2012). If the price of human capital were the key inefficiency in human capital markets, we would expect these policies to have no effect. We would expect an MVPF near zero since individuals would place no value on the additional information. Instead, the research suggests that these policies induce many students to change their schooling attainment, and our results suggest that these changes lead to very high MVPFs. This again points to informational constraints or other optimization frictions being important factors in determining educational investment. From an MVPF perspective, it suggests that perhaps policies relaxing these constraints have higher returns than subsidies operating through the tax schedule.

V.F Inverse Optimum: Inferring Social Preferences

Lastly, one can also use our results to infer the social preferences the informed these economic policy choices. This is referred to as the inverse optimum approach (e.g. Bourguignon and Spadaro, 2012; Zoutman et al., 2013, 2016,). The idea is that if one assumes a benchmark model in which policies (e.g. income tax schedule) are set optimally according to some social preferences, then one can use equation (3) to back out the implicit welfare weights placed on individuals in the population.

Returning to the 1993 tax reform example above, the social preferences that rationalize the tax schedule as an optimum require that one be roughly indifferent to providing $1.85 to top earners versus $1.12 to those on EITC. Put differently, the tax rates set by the 1993 OBRA tax change would be optimal if one had a social welfare function that places 1.65 (1.85/1.12) times as much weight on those at the bottom of the income distribution.

While we can rationalize the 1993 reform as the actions of a social-welfare-maximizing government, this is not possible when considering all of our MVPF estimates at once. Figure X adds policies targeting children in addition to the tax and transfer policies in Figure IXB. This reveals two inconsistencies with a social welfare function. First, when expanding to the broad set of policies, we find systematic evidence of the existence of Laffer effects for many policies targeting low-income children. These policies represent a Pareto improvement, and so we expect any welfare-maximizing government to have already taken advantage of this free lunch.[79]

Second, one can no longer draw an upward sloping line through the sets of policies in Figure X like one could potentially do in Figure IXA. From a theoretical perspective, the presence of

[79]The existence of policies with Laffer effects requires either that the government did not know about these policies until the time they were able to implement them, or that they did not know that the policies would pay for themselves.

very high MVPFs on transfers to low-income children and low MVPFs on transfers to adults are not well explained in an inverse optimum framework. Those patterns suggest that low-income children receive lower welfare weights than low- (and many high-) income adults. Such a result is particularly surprising given that income is often persistent across the lifecycle, meaning that low-income children are more likely to become low-income adults. In Appendix J we explore these implications in an intergenerational setting. We find that the observed pattern of MVPFs can only be justified if a government places very high welfare weights on low-income adults *who grew up wealthy*.[80] In other words, unless one has a particularly high affinity for the struggling children of the affluent, the results suggest that historical US government policies were not the implementation of an optimum.[81]

Variation over time One rationale for these high returns to investments in low-income children is lack of awareness. If this were the case, one might expect that the MVPFs of these policies targeting low-income children would have declined over time. Appendix Figure VI assesses this by plotting the child- and adult-average MVPFs separately by decade. We find no evidence for that pattern of decline. Instead, we find high MVPFs for policies targeting children throughout the past 50 years.[82] This finding is perhaps suggestive evidence that political constraints have prevented the government from enacting policies with high MVPFs over the long run.[83]

VI Lessons for Future Work

In this section, we discuss three implications for future economic research. First, we show how the MVPF framework facilitates a straightforward method to quantify the value of future work that reduces the statistical uncertainty in our estimates. Second, we show the value-added provided by measuring the MVPF relative to what is provided by a more traditional cost-benefit analysis. Third, we discuss how the intuitions of the MVPF framework might influence future empirical designs. The

[80] In Appendix J, we show that standard social welfare functions generally require (a) that MVPFs are on average the same for policies targeting different ages (society should be indifferent to *when* we provide welfare benefits) and (b) that if society places higher weight on poor as opposed to rich adults, it should also place greater weight on poor as opposed to rich children. The latter pattern follows because incomes are persistent across generations.

[81] Political constraints are a natural rationale for this deviation from optimal policy. We discuss this notion in more detail in the Conclusion below.

[82] The one exception to this pattern is the low average MVPF amongst child policies implemented in the 1970s. The child policies in that decade primarily consisted of job training programs that did not have significant effects on children's earnings.

[83] One potential constraint is that many of the policies we consider require spending by state and local governments, but much of the benefits accrue to the federal tax system. This perhaps diminishes local incentives for effective investment. We leave a formal analysis of this potential mechanism for future work.

key is to design experiments in a way that facilitates measuring willingness to pay. In particular, we discuss how 27 different welfare reform programs in the 1980s-90s randomized upwards of 100,000 participants into RCTs , but the nature of the research designs makes it infeasible to conduct reliable welfare analysis.

VI.A Value of Information in Evidence-Based Policy Making

Our MVPF estimates measure the welfare impact of a range of government policies. While it is our hope that these estimates can be useful for a policymaker seeking to conduct "evidence-based" policy, it is quite clear from Figure IVA that many of our individual policy estimates contain considerable sampling uncertainty. Here, we show how one can use the MVPF framework to understand the value of future research that reduces the uncertainty in our estimates. The MVPF framework provides a measure of the value of information because it is a price: it measures the price faced by the government to redistribute across beneficiaries of different types of policies. A welfare-maximizing government should be willing to pay to reduce the uncertainty in these prices, just as a consumer would be willing to pay to learn the true value of the products he or she buys.

There are many ways one could conceptualize reducing the various sources of modeling and sampling uncertainty in our estimates. In this section, we develop one simple approach to measure the value of reducing sampling uncertainty. (We defer an exhaustive treatment to future work.) We use this example to illustrate the value of future research that increases estimate precision, perhaps through improved access to larger administrative longitudinal datasets.[84]

Our conceptual experiment is organized as follows: suppose a policymaker is considering whether to raise revenue to spend an additional \$1 on policy j. The policy has a net cost to the government of G_j and a willingness to pay of WTP_j per dollar of programmatic cost. But, the policymaker does not know the true values of WTP_j and G_j. Instead, we assume she only observes the estimates, $\hat{WTP}_j$ and $\hat{G}_j$, and their sampling distributions.[85] We assume the policymaker has an uninformed prior about the impact of the policy so that the estimated sampling distribution reflects her belief about the policy's effects.

For simplicity, we assume the policy is financed with tax change that targets the same beneficiaries and has an MVPF of 1. A budget-neutral policy that increases taxes to spend on policy j

[84] The focus here is on reducing uncertainty amongst the observed outcomes of each program. Uncertainty regarding unobserved causal effects remains beyond the scope of this exercise.

[85] For simplicity, we assume programmatic costs are known and equal to their point estimates.

has a welfare gain of

$$U\left(WTP_j, G_j\right) = WTP_j - G_j$$

Ideally, the policymaker would wish to pursue this policy if and only if $U\left(WTP_j, G_j\right) > 0$ (i.e. the policy increases welfare). In practice, the policymaker only observes estimates and sampling distributions of these values. Utility is linear in WTP_j and G_j, so the policymaker will rely on the point estimates, $\hat{WTP_j}$ and $\hat{G_j}$, to make her decision. She will choose the policy if and only if $\hat{WTP_j} > \hat{G_j}$. The expected utility of this strategy is given by

$$\begin{aligned} EU^{Uninformed}(\hat{WTP_j}, \hat{G_j}) &= U(\hat{WTP_j}, \hat{G_j}) * 1\left\{U(\hat{WTP_j}, \hat{G_j}) > 0\right\} \\ &= \left(\hat{WTP_j} - \hat{G_j}\right) * 1\left\{\hat{WTP_j} > \hat{G_j}\right\} \end{aligned}$$

Now, suppose instead of spending \$1 on the policy, the policymaker can invest a fraction of this dollar, v_j, into learning more about the WTP_j and G_j of the policy before making this decision. We begin by considering a case where spending v_j allows the policymaker to perfectly learn WTP_j and G_j before deciding whether to invest in the policy. Once informed, the government chooses to pursue the policy if and only if $U\left(WTP_j, G_j\right) > 0$. Now it can decide to pursue the policy if and only if the true WTP exceeds the true costs. In that case, the net utility to the government is

$$U^{informed}\left(WTP_j, G_j, v_j\right) = \left(1 - v_j\right)\left(WTP_j - G_j\right) * 1\left\{WTP_j > G_j\right\} - v_j$$

where the first term is the surplus from investing the remaining fraction $1 - v_j$ in the policy and the second term is the cost of paying for the information.

The value to the government of learning the true willingness to pay and cost for policy j is the value of v_j^{info} which solves the following equation:

$$E\left[U^{informed}\left(WTP_j, FE_j, v_j^{info}\right) \mid \hat{WTP_j}, \hat{G_j}\right] = EU^{uninformed}(\hat{WTP_j}, \hat{G_j}) \tag{10}$$

Here, v_j^{info} equates the government's expected utility in the case where the it spends v_j^{info} to receive additional information and the case where it remains uninformed. The expectation in the LHS of equation (10) is taken with respect to the sampling distribution of the estimates. This implicitly defines v_j^{info} as the value that makes one indifferent to remaining uninformed versus paying for the information and making a decision based upon it.

Results We estimate the value of info in equation (10) both for each individual policy and for our category averages. Figure XI presents the results of v_j^{info} for each policy, j, plotted relative

to the age of the policy's beneficiaries. Broadly, we find the highest values of future research for policies with uncertain long-run impacts on children. For example, we estimate $v_{FS}^{info} = \$0.50$ for the introduction of food stamps. Moreover, we also find large values of information for policies with potential indirect policies on children and uncertain impacts on adults. We also find large values of information for college subsidies to parents (shown in green). This reflects the fact that these policies have highly uncertain impacts on college attainment, and small increases in attainment can translate into large gains. In contrast, we find smaller values of information for policies where the effects have been already precisely estimated. For example, we find the evidence-based policymaker would be willing to pay little to remove the statistical uncertainty in the estimated impact of disability insurance on labor earnings (e.g. we estimate the policymaker is willing to pay \$0 to learn the precise impact of assignment to a more lenient DI judge). This lower value of information reflects the relatively high precision of existing estimates in those studies.

Administrative vs. Survey Data: Long-Run Impacts of Food Stamps Our estimates in Figure XI report the value of learning the true impact of the policy. In practice, the true impact is never observable. That said, improved access to larger administrative datasets can help obtain more precise impacts of government policies. For example, a policymaker can decide whether a researcher should utilize a survey dataset for the analysis or obtain access to linked administrative data on the population.

To illustrate this decision, we consider the case of the introduction of food stamps discussed in Section III.C. Earlier work by Hoynes and Schanzenbach (2009) used the Panel Study of Income Dynamics (PSID) survey dataset to identify the long-run impact of food stamps on children's outcomes. More recently, Bailey et al. (2019) used linked Census data to estimate those effects more precisely. Here, we imagine that a policymaker is deciding whether to introduce food stamps based on the existing evidence. Consider the hypothetical example that they know the PSID estimates from Hoynes and Schanzenbach (2009), $\hat{WTP}^{PSID}$ and $\hat{FE}^{PSID}$. Suppose that they can instead invest v^{admin} to learn the estimates with the same statistical precision as those found in Bailey et al. (2019) based on Census data. Instead of learning the true value of WTP and FE, the policymaker learns $\hat{WTP}^{Census}$ and $\hat{FE}^{Census}$. The policymaker will expect these estimates to be drawn from the PSID sampling distribution but contain the standard errors found in the Census data estimates. The value of learning the Census estimates, v^{Census}, then solves

$$E\left[(1-v^{Census})\,U\left(\hat{WTP}^{Census},\hat{FE}^{Census}\right)1\left\{\hat{WTP}^{Census}>1-\hat{FE}^{Census}\right\}-v^{Census}\mid \hat{WTP}^{PSID},\hat{FE}^{PSID}\right] \quad (11)$$
$$=$$
$$U\left(\hat{WTP}^{PSID},\hat{FE}^{PSID}\right)1\left\{\hat{WTP}^{PSID}>1-\hat{FE}^{PSID}\right\}$$

The LHS of equation (11) is the expected value of investing in administrative data at a price v^{Census}. The RHS is the expected value of the policy if she makes her decision using the information in the PSID.

We reconstruct the estimates of the WTP and FE for the introduction of food stamps using the estimates from Hoynes and Schanzenbach (2009) in place of those in Bailey et al. (2019), normalizing by the mechanical program cost. This yields an infinite point estimate for our MVPF, and we find a willingness to pay estimate of 6.06 (95% CI of [-12.07, 23.78]) and cost of -0.19 (95% CI of [-5.19, 4.92]). These estimates are notably less precise than the estimates using the results from Bailey et al. (2019) that use Census data, which generate a WTP of 1.09 (95% CI of [-2.45, 4.55]).

Plugging these estimates into equation (11) suggests the policymaker would be willing to invest \$0.24 per dollar of investment in the food stamp program to learn the long-run estimates from Census data instead of PSID data. This exercise illustrates that if the policymaker only knew the PSID estimates, there would be a large value in learning additional information before making this investment decision.

This is, of course, a stylized exercise. We are imagining a policymaker that sees the ex-post evaluation of a policy prior to making her decision – something that is clearly not feasible. The goal here is merely to illustrate potential value of expanding access to administrative datasets that can generate more precise estimates of long-run policy impacts.

VI.B Comparison to BCR

While we focus on computing the MVPF for each policy, the most common form of welfare analysis in previous literature is cost-benefit analysis, as in equation (4). With that in mind, we compare our results to the benefit-cost ratios for the same policies. Figure XIIA plots the benefit cost ratio for a deadweight loss of $\phi = 50\%$ as in Heckman et al. (2010) as a function of the age of the beneficiary of the policy. Our general conclusion about the high returns to investment in children would remain true even if one used a cost-benefit ratio instead of the MVPF. The average cost-benefit ratio is 4.13 for child education, 5.30 for child health, and 6.78 for college policies. In contrast, we find smaller cost-benefit ratios for adult policies – often less than 1.

To directly compare the two methods of welfare analysis, Figure XIIB plots the BCR on the vertical axis (again for $\phi = 50\%$) against the MVPF on the horizontal axis. In general, we find a fairly monotonic relationship – policies with high BCRs also have high MVPFs. There are, however, some notable distinctions. For example, the Medicaid expansion to children born after September 30, 1983 has an infinite MVPF but a BCR of just 1.37. Similarly, the 1981 top tax rate reduction has an infinite MVPF but a cost benefit ratio of 1.67. By the standards of cost-benefit ratios these policies may not appear all that desirable, even though the MVPF point estimates imply that they pay for themselves and provide a Pareto improvement.

The difference between the MVPF and BCR in these cases reflects the fact that the benefit-cost ratio places all causal effects of the program in the numerator while the MVPF incorporates effects based on their incidence. In particular, the numerator of the MVPF captures the impacts on beneficiaries while the denominator captures all impacts on the government budget. In measuring the welfare effects of the 1983 Medicaid expansion and the 1981 tax cut, MVPF places all fiscal externalities in the denominator. The results show us that these policies have substantial benefits and limited or no net government cost. In the BCR framework these reforms would have been interpreted as high cost policies with substantial benefits.

The second crucial distinction between the MVPF and BCR is the way in which the two approaches conceptually close the budget constraint. While the MVPF closes the budget constraint by comparing MVPFs of different policies (and aggregating using Okun's bucket as in equation (3)), the BCR approach closes the budget constraint in a manner that ignores distributional incidence. The BCR approach simply imagines that the policy was funded by an increase in the marginal tax rate that led to a distortion in tax revenue and a DWL of ϕ. Consequently, the deadweight loss parameter ϕ in equation (4) is not context dependent and is instead universally applied across policies.

To see how this matters, consider the 1993 tax reform that simultaneously raised top marginal income tax rates and expanded the EITC. One could, in principle, use a BCR to evaluate whether the EITC expansion was desirable. As shown in Figure XIIA, the BCR for the 1993 EITC expansion is 0.74 after adjusting for a 50% DWL. The costs exceed the benefits and so, if the government were applying a strict cost-benefit test, we would not expect this policy to be implemented. This is because of the hypothetical 50% cost of raising the funds is too large to justify the expenditure.

That said, the goal of the EITC expansion was to provide redistributive benefits to low-income workers. Its MVPF is 1.12, near the highest amongst policies targeting adults. Rather than ruling

this out as a means of redistribution, we can compare the MVPF of the EITC to the MVPF of a tax increase used to fund this policy. Comparisons of MVPFs correspond to precise statements of social welfare using Okun's bucket. As noted above, the MVPF point estimate for the 1993 top tax rate change is 1.85. If society prefers giving $1.12 to a low-income worker on EITC to giving $1.85 to a high income individual facing the top marginal income tax rate, then the policy is welfare-enhancing despite its relatively low BCR.

VI.C Welfare Reform: Lessons for Future RCTs.

We end with a lesson of how an MVPF perspective can help inform the design of RCTs. Throughout, we aimed to include all possible MVPFs in the categories we considered. We included any policy where we thought we could provide reasonable measures of both costs and WTP. One set of notable omissions are the state-level welfare reforms made by states that sought to increase family self-sufficiency. Throughout the 1980s and early 1990s, states experimented with a range of reforms to cash welfare programs that imposed term limits, provided job training and other educational services, and provided job search and placement assistance.

The omission of these reforms is not because they were not analyzed. Many states rigorously evaluated the effect of these reforms. Upwards of 100,000 participants were enrolled into into 27 randomized controlled trials (RCTs) nationwide (Greenberg et al., 2010). These RCTs measured and provided a clear estimate of the net cost of each reform. However, the design of the policies enacted in each state makes it difficult to understand their welfare impacts. Generally, programs contained both a carrot and a stick. As a result, we cannot even accurately sign the WTP. As noted by MDRC who implemented the evaluations of these policies[86], "all [programs] contained a core quid pro quo arrangement in which the government would offer education, training, job search assistance, and support services to people receiving cash welfare, while most recipients — the majority of them single parents — would be required to participate in such services in order to qualify for benefits." While we can evaluate whether the government saved money, we do not know if the people in these programs benefitted from their participation. It may be that government revenue gains were the result of expanded job opportunities due to program participation. In that case, willingness to pay would be positive. By contrast, it may be that the government revenue gains were the result of stricter attendance requirements that drove individuals off welfare. In that

[86]See https://www.mdrc.org/project/evaluations-state-welfare-work-programs#design-site-data-sources, accessed on July 7, 2019.

case, willingness to pay would be negative.[87]

This highlights the value of isolating the carrot and the stick into separate RCTs. It also demonstrates the value of designing experiments to estimate individual WTP for non-market goods such as job training, job search assistance, or other educational policies. In Appendix Figure VII, we conduct a range of bounding exercises that attempt to construct lower and upper bounds on WTP for these welfare reform programs. Unfortunately, the bounds are very wide. In many cases, the policies are Pareto dominated, $MVPF < 0$, under one set of assumptions and represent a Pareto improvement, $MVPF = \infty$, under another set of assumptions.[88] Despite substantial expenditures on the evaluation of these reforms, the designs of these reforms in each state make it difficult to know whether this massive shift in the provision welfare benefits to low-income families led to an increase or decrease in welfare.

VII Conclusion

In this paper, we examine the Marginal Value of Public Funds of 133 different historical policies over the last half-century in the United States. We find a clear and persistent pattern that direct investments in children have yielded the largest MVPFs. There is a large "bang for the buck" associated with a range of expenditures on children from early education to child health insurance to college expenditures.

We also demonstrate that in a meaningful number of cases these policies pay for themselves. In particular, when government expenditures boost human capital, the resulting increase in net government revenue can offset the policy's upfront costs. From a taxpayer perspective, these expenditures on children are investments, rather than just transfers.

We find that opportunities for high return investments in children have persisted across policy categories for many decades. This is, however, no guarantee that all future investment in these categories will produce high MVPFs. Indeed, we find that MVPFs vary substantially within policy categories. Low-return policies exist even in high-return categories. This highlights the value of further understanding the mechanisms behind the high MVPFs of successful historical investments.

Even in cases where there is existing research, much still remains unknown about the welfare

[87] Previous work (Greenberg et al., 2010) has conducted a cost-benefit analysis of these reforms by assuming willingness to pay is given by after-tax earnings. However, if the term limit is what causes individuals to choose to move off of welfare and into the labor market (thus increasing earnings), the envelope theorem would suggest the WTP is negative, even if after-tax earnings increase.

[88] In fact, we find policies that follow this pattern in each sub-category of welfare reform programs. These sub-categories include job search assistance.

consequences of government policy. To that aim, we quantify the value of future work that uses new data to reduce estimate uncertainty. We show that in many cases, an evidence-based policymaker seeking to maximize social welfare should be willing to make substantial budgetary expenditures to learn more about policy effectiveness. In particular, our results highlight the value of expanded use of administrative data for policy analysis.

The 133 policies included in this paper are just a small subset of the policies that could be analyzed using the MVPF. We do not discuss the MVPF of crime policies, environmental policies, macroeconomic stabilization policies, or infrastructure policies, amongst many others. With careful tracking of willingness to pay and net costs, the MVPF can be used in any of these contexts. We leave that analysis for future work.

References

Abraham, K. G. and M. A. Clark (2006). Financial aid and students' college decisions: Evidence from the district of columbia tuition assistance grant program. *The Journal of Human Resources 41*(3), 578–610.

Ackerman, D., D. O. T. Treasury, J. Holtzblatt, C. B. Office, K. Masken, and I. R. Service (2009). The pattern of EITC claims over time: a panel data analysis. In *Conference Paper from IRS RC 2009: Internal Revenue Service Research Conference. Washington, DC: Department of the Treasury.*

Acton, R. (2018). The impact of public tuition subsidies on college enrollment decisions: Evidence from michigan.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics 84*(3), 488–500.

Akerlof, G. A. (1978). The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning. *The American economic review 68*(1), 8–19.

Almond, D., H. Hoynes, and D. W. Schanzenbach (2011). Inside the war on poverty: The impact of food stamps on birth outcomes. *The Review of Economics and Statistics 93*(2), 387–403.

American Institutes for Research (2017). Delta cost project. https://www.deltacostproject.org/delta-cost-project-database. Accessed April 26, 2019.

Andrews, I. and M. Kasy (Forthcoming). Identification of and correction for publication bias. *American Economic Review*.

Atkinson, A. B., T. Piketty, and E. Saez (2011). Top incomes in the long run of history. *Journal of economic literature 49*(1), 3–71.

Atkinson, A. B. and N. H. Stern (1974). Pigou, taxation and public goods. *The Review of Economic Studies 41*(1), 119–128.

Atkinson, A. B. and J. E. Stiglitz (1976). The design of tax structure: Direct versus indirect taxation. *Journal of public Economics 6*(1-2), 55–75.

Auerbach, A. (1985). The theory of excess burden and optimal taxation. In *Handbook of Public Economics*, Volume 1, pp. 61–127. Elsevier.

Auerbach, A. J. and J. R. Hines (2002). Taxation and economic efficiency. In *Handbook of Public Economics*, Volume 3, pp. 1347–1421. Elsevier.

Auten, G. and R. Carroll (1999). The effect of income taxes on household income. *Review of economics and statistics 81*(4), 681–693.

Autor, D. H., M. Duggan, K. Greenberg, and D. S. Lyle (2016). The impact of disability benefits on labor supply: Evidence from the VA's disability compensation program. *American Economic Journal: Applied Economics 8*(3), 31–68.

Bailey, M., H. Hoynes, M. Rossin-Slater, and R. Walker (2019). Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence from the Food Stamps Program. Goldman School of Public Policy Working Paper.

Baily, M. (1978). Some aspects of optimal unemployment insurance. *Journal of Public Economics 10*, 379–402.

Barnett, W. and L. N. Masse (2007). Comparative benefit cost analysis of the abecedarian program and its policy implications. *Economics of Education Review 26*(1), 113 – 125.

Bartik, T. J., B. Hershbein, and M. Lachowska (2016). The merits of universal scholarships: Benefit-cost evidence from the kalamazoo promise. *Journal of Benefit-Cost Analysis 7*(3), 400–433.

Bartik, T. J., B. J. Hershbein, and M. Lachowska (2017). The effects of the kalamazoo promise scholarship on college enrollment, persistence, and completion.

Bastian, J. and K. Michelmore (2018). The long-term impact of the earned income tax credit on children's education and employment outcomes. *Journal of Labor Economics 36*(4), 1127–1163.

Bettinger, E. (2004). How financial aid affects persistence. In *College choices: The economics of where to go, when to go, and how to pay for it*, pp. 207–238. University of Chicago Press.

Bettinger, E., O. Gurantz, L. Kawano, B. Sacerdote, and M. Stevens (2019). The long-run impacts of financial aid: Evidence from california's cal grant. *American Economic Journal: Economic Policy 11*(1), 64–94.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2012). The role of application assistance and information in college decisions: Results from the h&r block fafsa experiment. *The Quarterly Journal of Economics 127*(3), 1205–1242.

Bhargava, S. and D. Manoli (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an irs field experiment. *American Economic Review 105*(11), 3489–3529.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2007). From the cradle to the labor market? The effect of birth weight on adult outcomes. *The Quarterly Journal of Economics 122*(1), 409–439.

Blank, R. M. and P. Ruggles (1996). When Do Women Use Aid to Families with Dependent Children and Food Stamps? The Dynamics of Eligibility Versus Participation. *The Journal of Human Resources 31*(1), 57–89.

Bloom, D., J. J. Kemple, P. Morris, S. Scrivener, N. Verma, R. Hendra, D. Adams-Ciardullo, D. Seith, and J. Walter (2000). Final report on florida's initial time-limited welfare program. *New York: Manpower Demonstration Research Corporation*.

Bloom, D., S. Scrivener, C. Michalopoulos, P. Morris, R. Hendra, D. Adams-Ciardullo, and J. Walter (2002). Jobs first: Final report on connecticut's welfare reform initiative. *ERIC*.

Bloom, H. S., L. L. Orr, S. H. Bell, G. Cave, F. Doolittle, W. Lin, and J. M. Bos (1997). The Benefits and Costs of JTPA Title II-A Programs: Key Findings from the National Job Training Partnership Act Study. *The Journal of Human Resources 32*(3), 549–576.

Bloom, H. S., J. A. Riccio, and N. Verma (2005). Promoting Work In Public Housing: The Effectiveness of Jobs-Plus. *New York: Manpower Demonstration Research Corporation*.

Boudreaux, M. H., E. Golberstein, and D. D. McAlpine (2016). The long-term impacts of medicaid exposure in early childhood: Evidence from the program's origin. *Journal of Health Economics 45*, 161–175.

Bourguignon, F. and A. Spadaro (2012). Tax–benefit revealed social preferences. *Journal of Economic Inequality 10*(1), 75–108.

Bovenberg, A. L. and B. Jacobs (2005). Redistribution and education subsidies are siamese twins. *Journal of Public Economics 89*(11-12).

Brock, T. et al. (1993). Unpaid work experience for welfare recipients: Findings and lessons from mdrc research. mdrc working papers. *New York: Manpower Demonstration Research Corporation*.

Brown, D., A. E. Kowalski, and I. Z. Lurie (2017). Long-term impacts of childhood medicaid expansions on outcomes in adulthood. Working Paper 20835, National Bureau of Economic Research.

Brown, D. W., A. E. Kowalski, and I. Z. Lurie (2015). Medicaid as an investment in children: What is the long-term impact on tax receipts? Working Paper 20835, National Bureau of Economic Research.

Bruce, D. J. and C. K. Carruthers (2014). Jackpot? the impact of lottery scholarships on enrollment in tennessee. *Journal of Urban Economics 81*, 30–44.

Bulman, G. B. and C. M. Hoxby (2015). The returns to the federal tax credits for higher education. *Tax Policy and the Economy 29*(1), 13–88.

Cabral, M. and N. Mahoney (2019). Externalities and taxation of supplemental insurance: A study of medicare and medigap. *American Economic Journal: Applied Economics 11*(2), 37–73.

Camerer, C. F., A. Dreber, E. Forsell, T.-H. Ho, J. Huber, M. Johannesson, M. Kirchler, J. Almenberg, A. Altmejd, T. Chan, et al. (2016). Evaluating replicability of laboratory experiments in economics. *Science 351*(6280), 1433–1436.

Card, D., A. Johnston, P. Leung, A. Mas, and Z. Pei (2015). The effect of unemployment benefits on the duration of unemployment insurance receipt: New evidence from a regression kink design in missouri, 2003-2013. *American Economic Review: Papers and Proceedings 105*(5), 126–130.

Carroll, R. (1998). Do taxpayers really respond to changes in tax rates? Evidence from the 1993 tax act. Working Paper 79, US Department of the Treasury Office of Tax Analysis.

Carruthers, C. K. and J. G. Welch (2019). Not whether, but where? pell grants and college choices. *Journal of Public Economics 172*, 1–19.

Castleman, B. L. and B. T. Long (2016). Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. *Journal of Labor Economics 34*(4), 1023–1073.

Cave, G., H. Bos, F. Doolittle, and C. Toussaint (1993). Jobstart: Final Report on a Program for School Dropouts. *New York: Manpower Demonstration Research Corporation*.

Chetty, R. (2006). A general formula for the optimal level of social insurance. *Journal of Public Economics 90*, 1879–1901.

Chetty, R. (2008). Moral hazard versus liquidity and optimal unemployment insurance. *Journal of Political Economy 116*(2), 173–234.

Chetty, R., N. Hendren, and L. F. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *The American Economic Review 106*(4), 855–902.

Cohodes, S. R. and J. S. Goodman (2014). Merit aid, college quality, and college completion: Massachusetts' adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics 6*(4), 251–85.

Cornwell, C., K. H. Lee, and D. Mustard (2003). The effects of merit-based financial aid on course enrollment, withdrawal and completion in college. Working paper, Institute for the Study of Labor (IZA).

Cornwell, C., D. Mustard, and D. Sridhar (2006). The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE Scholarship. *Journal of Labor Economics*.

Couch, K. A. (1992). New evidence on the long-term effects of employment training programs. *Journal of Labor Economics 10*(4), 380–388.

Currie, J. (1994). Welfare and the well-being of children: The relative effectiveness of cash and in-kind transfers. *Tax policy and the economy 8*, 1–43.

Currie, J. and N. Cole (1993). Welfare and child health: The link between afdc participation and birth weight. *The American Economic Review 83*(4), 971–985.

Currie, J. and F. Gahvari (2008). Transfers in cash and in-kind: Theory meets the data. *Journal of Economic Literature 46*(2), 333–383.

Currie, J. and J. Gruber (1996). Saving babies: the efficacy and cost of recent changes in the medicaid eligibility of pregnant women. *Journal of political Economy 104*(6), 1263–1296.

Currie, J. and E. Moretti (2006). Did the introduction of food stamps affect birth outcomes in california? Working paper 06-20, National Poverty Center.

Cutler, D. M. and J. Gruber (1996). Does public insurance crowd out private insurance? *The Quarterly Journal of Economics 111*(2), 391–430.

Dave, D., S. Decker, R. Kaestner, and K. Simon (2015). The Effect of Medicaid Expansions in the Late 1980s and Early 1990s on the Labor Supply of Pregnant Women. *American Journal of Health Economics 1*(2), 195–193.

DeLong, J. B., L. H. Summers, M. Feldstein, and V. A. Ramey (2012). Fiscal policy in a depressed economy. *Brookings Papers on Economic Activity*, 233–297.

Deming, D. J. and C. R. Walters (2017). The impact of price caps and spending cuts on u.s. postsecondary attainment. Working Paper 23736, National Bureau of Economic Research.

Denning, J. T. (2017). College on the cheap: Consequences of community college tuition reductions. *American Economic Journal: Economic Policy 9*(2), 155–88.

Denning, J. T., B. M. Marx, and L. J. Turner (2019). Propelled: The effects of grants on graduation, earnings, and welfare. *American Economic Journal: Applied Economics 11*(3), 193–224.

Deshpande, M. (2016). Does welfare inhibit success? the long-term effects of removing low-income youth from the disability rolls. *American Economic Review 106*(11), 3300–3330.

Diamond, P. and E. Saez (2011). The case for a progressive tax: from basic research to policy recommendations. *Journal of Economic Perspectives 25*(4), 165–90.

Dynarski, S. (2000). Hope for whom? financial aid for the middle class and its impact on college attendance. *National Tax Journal 53*(3), 629–62.

Dynarski, S., C. Libassi, K. Michelmore, and S. Owen (2018). Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students. Working Paper 25349, National Bureau of Economic Research.

Dynarski, S. M. (2003). Does aid matter? measuring the effect of student aid on college attendance and completion. *American Economic Review 93*(1), 279–288.

East, C. N. (2018). Immigrants' labor supply response to food stamp access. *Labour Economics 51*, 202–226.

Efron, B. (1982). *The jackknife, the bootstrap, and other resampling plans*, Volume 38. Siam.

Eissa, N. and H. W. Hoynes (2004). Taxes and the labor market participation of married couples: the earned income tax credit. *Journal of public Economics 88*(9-10), 1931–1958.

Eissa, N., H. J. Kleven, and C. T. Kreiner (2008). Evaluation of four tax reforms in the united states: Labor supply and welfare effects for single mothers. *Journal of Public Economics 92*(3-4), 795–816.

Eissa, N. and J. Liebman (1996). Labor supply responses to the earned income tax credit. *The Quarterly Journal of Economics 111*(2), 605–637.

Fein, D. and J. Hamadyk (2018). Bridging the opportunity divide for low-income youth: Implementation and early impacts of the year up program. OPRE Report 2018-65.

Finkelstein, A., N. Hendren, and E. F. Luttmer (Forthcoming). The Value of Medicaid: Interpreting Results from the Oregon Health Insurance Experiment. *Journal of Political Economy*.

Finkelstein, A., N. Hendren, and M. Shepard (2019). Subsidizing health insurance for low-income adults: Evidence from massachusetts. *American Economic Review 109*(4), 1530–1567.

Finkelstein, A. and R. McKnight (2008). What did medicare do? the initial impact of medicare on mortality and out of pocket medical spending. *Journal of public economics 92*(7), 1644–1668.

Finkelstein, A. and M. J. Notowidigdo (2019, 05). Take-Up and Targeting: Experimental Evidence from SNAP*. *The Quarterly Journal of Economics 134*(3), 1505–1556.

Freedman, S., D. Friedlander, W. Lin, and A. Schweder (1996). The gain evaluation: Five-year impacts on employment, earnings and afdc receipt. *New York: Manpower Demonstration Research Corporation*.

Freedman, S., J. T. Knab, L. A. Gennetian, and D. Navarro (2000). The los angeles jobs-first gain evaluation: Final report on a work first program in a major urban center. *New York: Manpower Demonstration Research Corporation*.

French, E. and J. Song (2014). The effect of disability insurance receipt on labor supply. *American Economic Journal: Economic Policy 6*(2), 291–337.

Ganong, P. and P. J. Noel (2019). Consumer spending during unemployment: Positive and normative implications. *American Economic Review 109*(7), 2383–2424.

García, J. L., J. J. Heckman, D. E. Leaf, and M. J. Prados (2017). Quantifying the life-cycle benefits of a prototypical early childhood program. Working Paper 23479, National Bureau of Economic Research.

Gelber, A., T. J. Moore, and A. Strand (2017). The effect of disability insurance payments on beneficiaries' earnings. *American Economic Journal: Economic Policy 9*(3), 229–61.

Goering, J., J. Kraft, J. Feins, D. McInnis, M. J. Holin, and H. Elhassan (1999). Moving to opportunity for fair housing demonstration program: Current status and initial findings. *Washington, DC: US Department of Housing and Urban Development*.

Goldrick-Rab, S., R. Kelchen, D. N. Harris, and J. Benson (2016). Reducing Income Inequality in Educational Attainment: Experimental Evidence on the Impact of Financial Aid on College Completion. *American Journal of Sociology 121*(6), 1762–1817.

Golosov, M. and A. Tsyvinski (2007). Optimal taxation with endogenous insurance markets. *The Quarterly Journal of Economics 122*(2), 487–534.

Goodman, J. (2008). Who merits financial aid?: Massachusetts' adams scholarship. *Journal of public Economics 92*(10-11), 2121–2131.

Goodman-Bacon, A. (2017). The long-run effects of childhood insurance coverage: Medicaid implementation, adult health, and labor market outcomes. Working Paper 22899, National Bureau of Economic Research.

Greenberg, D. H., V. Deitch, and G. Hamilton (2010). A synthesis of random assignment benefit-cost studies of welfare-to-work programs. *Journal of Benefit-Cost Analysis 1*(1), 1–30.

Grogger, J. (2003). The effects of time limits, the eitc, and other policy changes on welfare use, work, and income among female-headed families. *Review of Economics and statistics 85*(2), 394–408.

Gruber, J. (1997). The Consumption Smoothing Benefits of Unemployment Insurance. *American Economic Review 87*(1), 192–205.

Hamilton, G., S. Freedman, L. Gennetian, C. Michalopoulos, J. Walter, D. Adams-Ciardullo, A. Gassman-Pines, S. McGroder, M. Zaslow, J. Brooks, S. Ahluwalia, E. Small, and B. Ricchetti (2001). How Effective Are Different Welfare-to-Work Approaches? Five-Year Adult and Child Impacts for Eleven Programs. National Evaluation of Welfare-to-Work Strategies. *New York: Manpower Demonstration Research Corporation*.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science 312*(5782), 1900–1902.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010). The rate of return to the highscope perry preschool program. *Journal of Public Economics 94*(1-2), 114–128.

Heckman, J. J., R. Pinto, A. M. Shaikh, and A. Yavitz (2011). Inference with imperfect randomization: The case of the perry preschool program. Working Paper 16935, National Bureau of Economic Research.

Heim, B. T. (2009). The effect of recent tax changes on taxable income: Evidence from a new panel of tax returns. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management 28*(1), 147–163.

Helburn, S. W. (1995). Cost, Quality and Child Outcomes in Child Care Centers. Technical Report, Public Report, and Executive Summary. *New York: Manpower Demonstration Research Corporation*.

Hendra, R., D. H. Greenberg, G. Hamilton, A. Oppenheim, A. Pennington, K. Schaberg, and B. L. Tessler (2016). Encouraging evidence on a sector-focused advancement strategy: Two-year impacts from the workadvance demonstration. *New York: Manpower Demonstration Research Corporation*.

Hendren, N. (2016). The policy elasticity. *Tax Policy and the Economy 30*(1), 51–89.

Hendren, N. (2017a). Efficient welfare weights. Working Paper 20351, National Bureau of Economic Research.

Hendren, N. (2017b). Knowledge of future job loss and implications for unemployment insurance. *American Economic Review 107*(7), 1778–1823.

Hendren, N. (2017c). Measuring ex-ante welfare in insurance markets. Working Paper 24470, National Bureau of Economic Research.

Hollister, R. G., P. Kemper, and R. A. Maynard (1984). The national supported work demonstration.

Hotz, J. and K. Scholz (2003). The earned income tax credit. In *Means-tested transfer programs in the United States*, pp. 141–198. University of Chicago press.

Hoxby, C. M. and G. B. Bulman (2016). The effects of the tax deduction for postsecondary tuition: Implications for structuring tax-based aid. *Economics of Education Review 51*, 23–60.

Hoynes, H., M. Page, and A. H. Stevens (2011). Can targeted transfers improve birth outcomes?: Evidence from the introduction of the WIC program. *Journal of Public Economics 95*(7-8), 813–827.

Hoynes, H., D. W. Schanzenbach, and D. Almond (2016). Long-run impacts of childhood access to the safety net. *American Economic Review 106*(4), 903–34.

Hoynes, H. W. and A. J. Patel (2018). Effective policy for reducing poverty and inequality? the earned income tax credit and the distribution of income. *Journal of Human Resources 53*(4), 859–890.

Hoynes, H. W. and D. W. Schanzenbach (2009). Consumption responses to in-kind transfers: Evidence from the introduction of the food stamp program. *American Economic Journal: Economic Policy 1*(4), 109–139.

Hoynes, H. W. and D. W. Schanzenbach (2012). Work incentives and the food stamp program. *Journal of Public Economics 96*(1-2), 151–162.

Hoynes, H. W. and D. W. Schanzenbach (2018). Safety net investments in children. *Brookings Papers on Economic Activity*, 89–133.

Hylland, A. and R. Zeckhauser (1981). Distributional objectives should affect taxes but not program choice or design. In *Measurement in Public Choice*, pp. 123–143. Springer.

Hyman, J. (2017). Does money matter in the long run? effects of school spending on educational attainment. *American Economic Journal: Economic Policy 9*(4), 256–80.

Jackson, C. K., C. Persico, and R. C. Johnson (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics 131*(1), 157–218.

Jacob, B. A. and J. Ludwig (2012). The effects of housing assistance on labor supply: Evidence from a voucher lottery. *American Economic Review 102*(1), 272–304.

Jacob, B. A., J. Ludwig, and M. Kapustin (2014). The impact of housing assistance on child outcomes: Evidence from a randomized housing lottery. *The Quarterly Journal of Economics 130*(1), 465–506.

Johnston, A. C. and A. Mas (2018). Potential unemployment insurance duration and labor supply: The individual and market-level response to a benefit cut. *Journal of Political Economy 126*(6), 2480–2522.

Jones, D. and I. Marinescu (2018). The labor market impacts of universal and permanent cash transfers: Evidence from the alaska permanent fund. Working Paper 24312, National Bureau of Economic Research.

Kane, T. J. (1994). College entry by blacks since 1970: The role of college costs, family background, and the returns to education. *Journal of political Economy 102*(5), 878–911.

Katz, L. F. and B. D. Meyer (1990). The impact of the potential duration of unemployment benefits on the duration of unemployment. *Journal of Public Economics 41*(1), 45–72.

Kawano, L., C. Weber, and A. Whitten (2016). Estimating the elasticity of broad income for high-income taxpayers. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2852048, Accessed 7/7/19.

Kemple, J. J. et al. (1995). Florida's Project Independence. Benefits, Costs, and Two-Year Impacts of Florida's JOBS Program.

Kleven, H. J. and C. T. Kreiner (2006). The marginal cost of public funds: Hours of work versus labor force participation. *Journal of Public Economics 90*(10-11), 1955–1973.

Kline, P. and C. R. Walters (2016, 07). Evaluating Public Programs with Close Substitutes: The Case of Head Start. *The Quarterly Journal of Economics 131*(4), 1795–1848.

Kroft, K. and M. J. Notowidigdo (2016). Should unemployment insurance vary with the unemployment rate? theory and evidence. *Review of Economic Studies 83*(3), 1092–1124.

LaLumia, S. (2012). Tax preferences for higher education and adult college enrollment. *National Tax Journal 65*(1), 59–92.

Landais, C. (2015). Assessing the welfare effects of unemployment benefits using the regression kink design. *American Economic Journal: Economic Policy 7*(4), 243–278.

Lindner, S. (2016). How do unemployment insurance benefits affect the decision to apply for social security disability insurance? *Journal of Human Resources 51*(1), 62–94.

Lo Sasso, A. T. and D. G. Seamster (2007). How federal and state policies affected hospital uncompensated care provision in the 1990s. *Medical Care Research and Review 64*(6), 731–744.

Long, B. (2004). *College Choices: The Economics of where to go, when to go, and how to pay for it*, Chapter 3, pp. 101–168. The University of Chicago Press Books.

Ludwig, J. and D. L. Miller (2007). Does head start improve children's life chances? evidence from a regression discontinuity design. *The Quarterly journal of economics 122*(1), 159–208.

Maestas, N., K. J. Mullen, and A. Strand (2013). Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt. *American Economic Review 103*(5), 1797–1829.

Marx, B. M. and L. J. Turner (2018). Borrowing trouble? human capital investment with opt-in costs and implications for the effectiveness of grant aid. *American Economic Journal: Applied Economics 10*(2), 163–201.

Masse, L. N. and W. S. Barnett (2002). A benefit cost analysis of the abecedarian early childhood intervention.

Mayshar, J. (1990). On measures of excess burden and their applications. *Journal of Public Economics 43*(3), 263–89.

Meyer, B. D. and W. K. C. Mok (2007). Quasi-experimental evidence on the effects of unemployment insurance from new york state. Working Paper 12865, National Bureau of Economic Research.

Meyer, B. D. and D. T. Rosenbaum (2001). Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers. *The Quarterly Journal of Economics 116*(3), 1063–1114.

Miller, C., L. F. Katz, G. Azurdia, A. Isen, and C. B. Schultz (2017). Expanding the earned income tax credit for workers without dependent children: Interim findings from the paycheck plus demonstration in new york city. *New York: MDRC*.

Miller, C., V. Knox, L. A. Gennetian, M. Dodoo, J. A. Hunter, and C. Redcross (2000). Reforming Welfare and Rewarding Work: Final Report on the Minnesota Family Investment Program. Volume 1: Effects on Adults and Volume 2: Effects on Children.

Miller, S. and L. R. Wherry (2018). The long-term effects of early life medicaid coverage. *Journal of Human Resources*.

Mills, G., D. Gubits, L. Orr, D. Long, J. Feins, B. Kaul, M. Wood, A. Jones, et al. (2006). Effects of housing vouchers on welfare families. *Washington, DC: US Department of Housing and Urban Development, Office of Policy Development and Research. 8*, 2010.

Mirrlees, J. (1976). Optimal tax theory: A synthesis. *Journal of Public Economics 6*(4), 327–358.

Mirrlees, J. A. (1971). An exploration into the theory of optimal income taxation. *The Review of Economic Studies 38*(2), 175–208.

Moffitt, R. A. (2002). Welfare programs and labor supply. *Handbook of public economics 4*, 2393–2430.

Moffitt, R. A. (2003). *Means-tested transfer programs in the United States*. University of Chicago Press Chicago:.

Mueller, A. I., J. Rothstein, and T. M. von Wachter (2016). Unemployment insurance and disability insurance in the great recession. *Journal of Labor Economics 34*(S1), S445–S475.

Okun, A. (1975). *Equality and Efficiency*. Brookings Institution Press.

Price, D. J. and J. Song (2018). The long-term effects of cash assistance. Working paper.

Rea, D. and T. Burton (2018). New evidence on the heckman curve.

Reynolds, A. J., J. A. Temple, D. L. Robertson, and E. A. Mann (2002). Age 21 Cost-Benefit Analysis of the Title I Chicago Child-Parent Centers. *Educational Evaluation and Policy Analysis 24*(4), 267–303.

Riccio, J. A. (2006). Jobs–plus: A promising strategy for increasing employment and self–sufficiency among public housing residents. Technical report, Presented Before the Subcommittee on Federalism and the Census, House Committee on Government Reform.

Saez, E. (2003). The effect of marginal tax rates on income: a panel study of 'bracket creep'. *Journal of Public Economics 87*(5-6), 1231–1258.

Saez, E., J. Slemrod, and S. H. Giertz (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature 50*(1), 3–50.

Sanbonmatsu, L., L. F. Katz, J. Ludwig, L. A. Gennetian, G. J. Duncan, R. C. Kessler, E. K. Adam, T. McDade, and S. T. Lindau (2011). Moving to opportunity for fair housing demonstration program: Final impacts evaluation. US Dept of Housing and Urban Development.

Schaberg, K. (2017). Can sector strategies promote longer-term effects? three-year impacts from the workadvance demonstration. *New York: Manpower Demonstration Research Corporation*.

Schmieder, J. F. and T. Von Wachter (2016). The effects of unemployment insurance benefits: New evidence and interpretation. *Annual Review of Economics 106*(3), 547–581.

Schochet, P. Z. (2018). National job corps study: 20-year follow-up study using tax data. *Mathematica Policy Research Report*.

Schochet, P. Z., J. Burghardt, and S. McConnell (2008). Does job corps work? impact findings from the national job corps study. *The American Economic Review 98*(5), 1864–1886.

Schochet, P. Z., J. A. Burghardt, S. M. McConnell, et al. (2006). National job corps study and longer-term follow-up study: impact and benefit-cost findings using survey and summary earnings records data. *US Department of Labor, Employment and Training Administration*.

Scholz, J. K. (1993). The earned income tax credit: Participation, compliance, and antipoverty effectiveness. Institute for Research on Poverty Discussion Papers 1020-93.

Scrivener, S., R. Hendra, C. Redcross, D. Bloom, C. Michalopoulos, and J. Walter (2002). WRP: Final Report on Vermont's Welfare Restructuring Project. *New York: Manpower Demonstration Research Corporation*.

Seftor, N. S. and S. E. Turner (2002). Back to school: Federal student aid policy and adult college enrollment. *Journal of Human resources*, 336–352.

Slemrod, J. and S. Yitzhaki (1996). The social cost of taxation and the marginal cost of funds. *International Monetary Fund Staff Papers 43*(1), 172–98.

Slemrod, J. and S. Yitzhaki (2001). Integrating expenditure and tax decisions: The marginal cost of funds and the marginal benefit of projects. *National Tax Journal 54*(2), 189–202.

Solon, G. (1985). Work incentive effects of taxing unemployment benefits. *Econometrica 53*(2), 295–306.

Stantcheva, S. (2017). Optimal taxation and human capital policies over the life cycle. *Journal of Political Economy 125*(6), 1931–1990.

Stiglitz, J. E. and P. Dasgupta (1971, April). Differential taxation, public goods, and economic efficiency. *The Review of Economic Studies 38*(2), 151–174.

Turner, N. (2011). The effect of tax-based federal student aid on college enrollment. *National Tax Journal 64*(3), 839–61.

United States Census Bureau (1966). Current population reports: Consumer income. https://www2.census.gov/prod2/popscan/p60-049.pdf. Accessed 7/7/19.

U.S. Department of Education Office of Postsecondary Education (2010). 2009-2010 Federal Pell Grant Program End-of-Year Report. https://www2.ed.gov/finaid/prof/resources/data/pell-2009-10/pell-eoy-09-10.pdf. Accessed 7/7/19.

Von Wachter, T., J. Song, and J. Manchester (2011). Trends in employment and earnings of allowed and rejected applicants to the social security disability insurance program. *American economic review 101*(7), 3308–29.

Weimer, D. (2009). *Cost-benefit analysis and public policy*, Volume 1. John Wiley & Sons.

Weinzierl, M. (2011). The surprising power of age-dependent taxes. *The Review of Economic Studies 78*(4), 1490–1518.

Werning, I. (2007). Optimal fiscal policy with redistribution. *The Quarterly Journal of Economics 122*(3), 925–967.

Wherry, L. R. and B. D. Meyer (2016). Saving teens: using a policy discontinuity to estimate the effects of medicaid eligibility. *Journal of Human Resources 51*(3), 556–588.

Wherry, L. R., S. Miller, R. Kaestner, and B. D. Meyer (2018). Childhood medicaid coverage and later-life health care utilization. *Review of Economics and Statistics 100*(2), 287–302.

Whitmore, D. (2002). What are food stamps worth? *Princeton University Industrial Relations Section Working Paper #468*.

Wood, M., J. Turnham, and G. Mills (2008). Housing affordability and family well-being: Results from the housing voucher evaluation. *Housing Policy Debate 19*(2), 367–412.

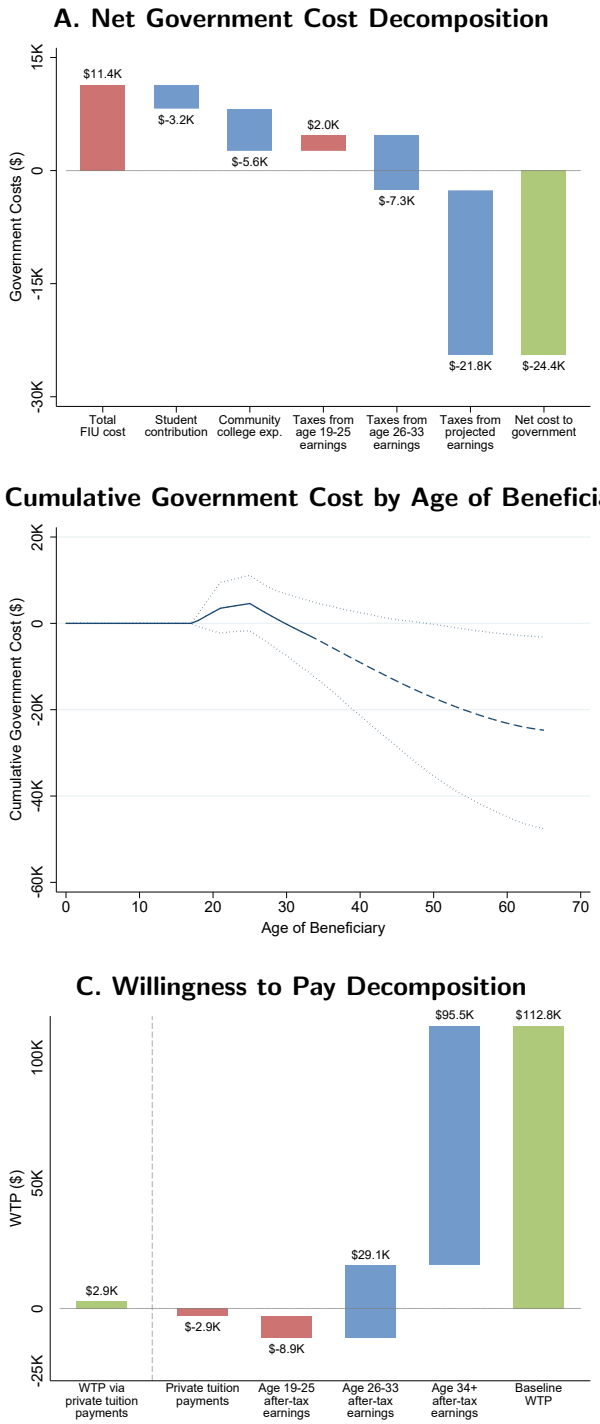
WSIPP (2018). Benefit-cost technical documentation. Technical report, Washington State Institute for Public Policy.

Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics 32*(4), 711–754.

Zoutman, F. T., B. Jacobs, and E. Jongen (2013). Optimal Redistributive Taxes and Redistributive Preferences in the Netherlands. Unpublished, available at https://personal.eur.nl/bjacobs/OTP.pdf, accessed 7/7/19.

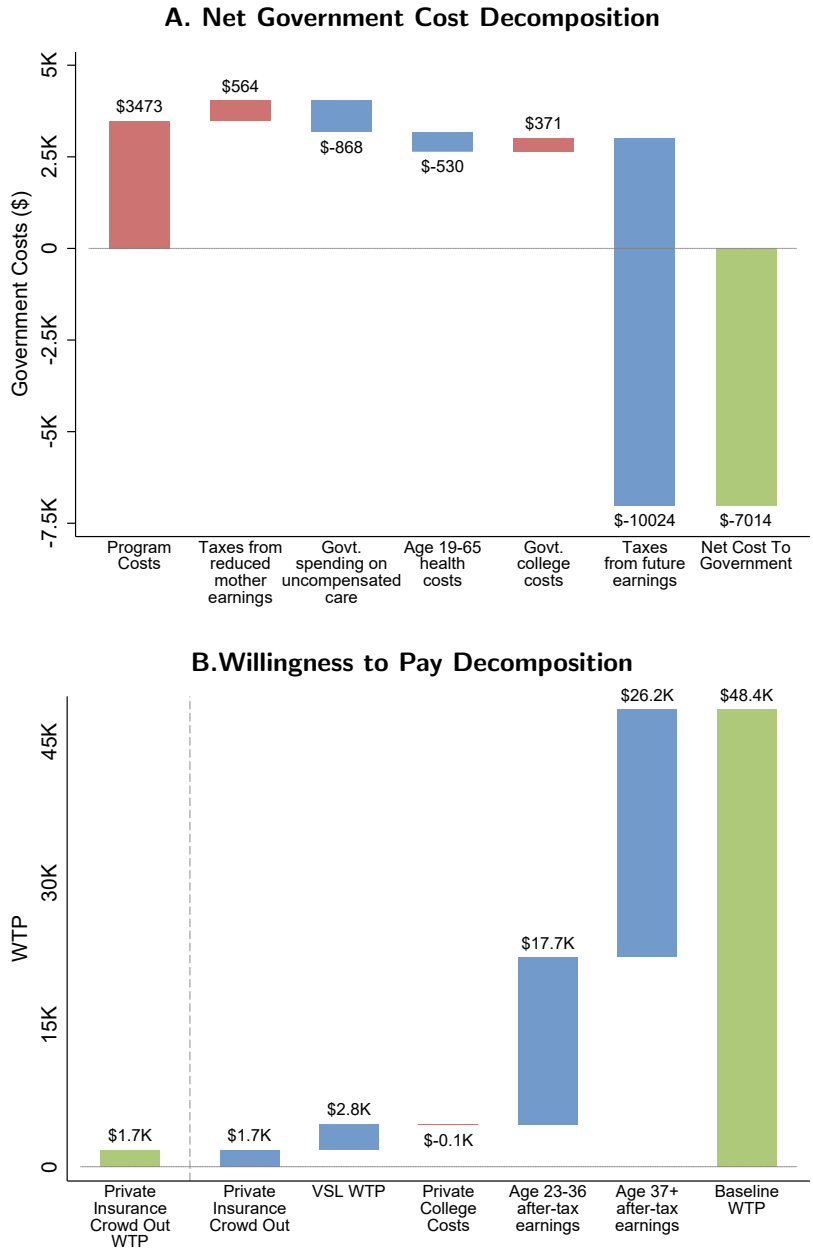
Zoutman, F. T., B. Jacobs, and E. Jongen (2016). Redistributive politics and the tyranny of the middle class. Unpublished, available at https://dx.doi.org/10.2139/ssrn.2769798, accessed 7/7/19.

FIGURE I: WTP and Cost Components for Admission to Florida International University



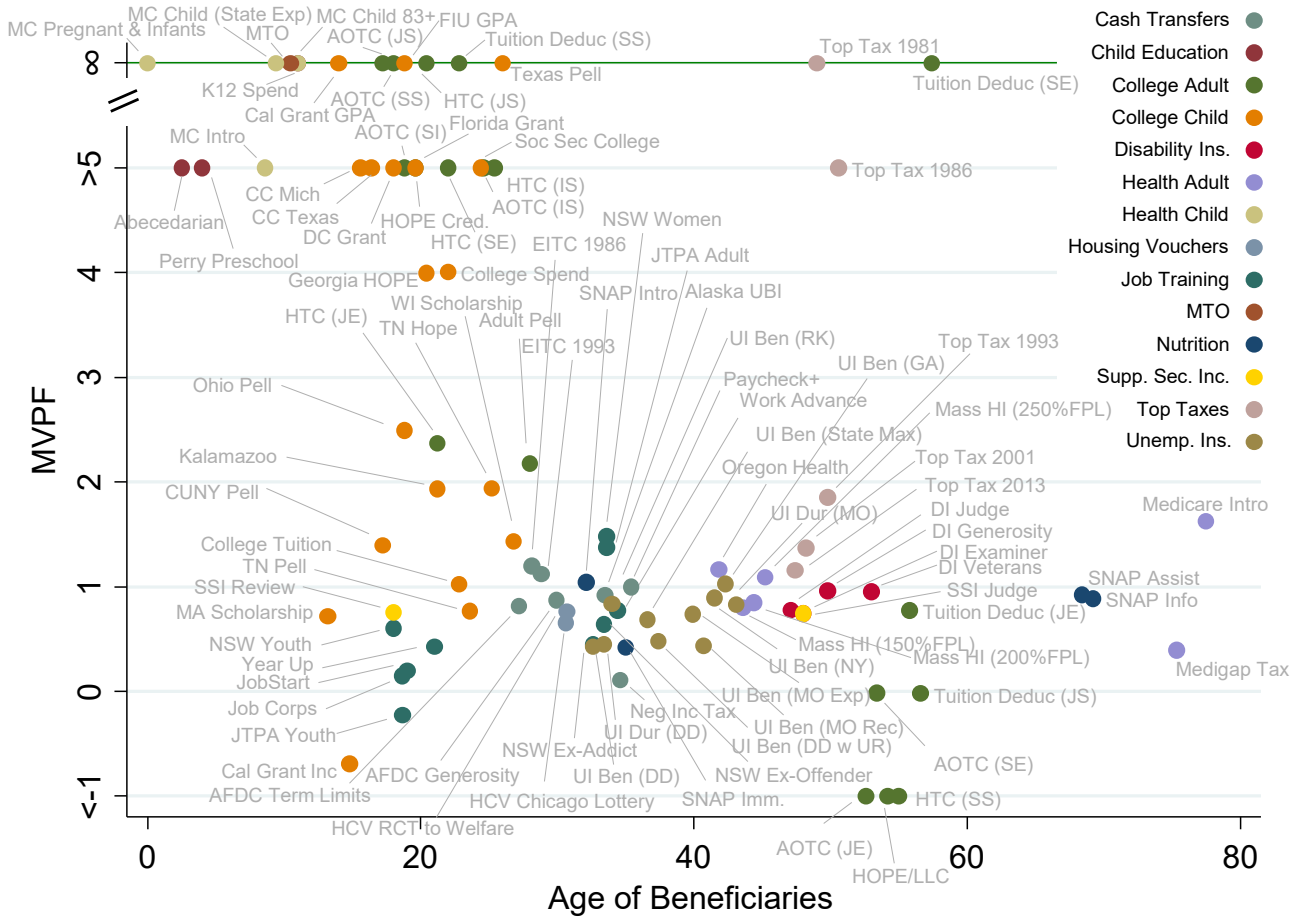
Notes: This figure illustrates the cost and willingness to pay components for admission to Florida International University as studied in Zimmerman (2014). Panel A breaks the total cost down into its various components, including increased student payments on tuition, reduced government spending on community colleges, and the changes in tax revenue from earnings. Panel B shows cumulative discounted cost of the policy over the lifetime of the beneficiary. The solid line represents cumulative costs for ages up until 33, the oldest age at which incomes are observed in Zimmerman (2014). The dashed line shows total costs inclusive of projected costs at subsequent ages. The projection method is detailed in Section III.1 and in Appendix I. Panel C reports the components of our WTP calculations. The conservative WTP is given by the tuition payments. The point estimate measures WTP as the change in incomes after taxes and expenses on tuition. The dotted lines provide the 95% bootstrap (pointwise) confidence intervals with adjustments discussed in Appendix H. All numbers in 2005 dollars deflating using the CPI-U-RS and discounted using a 3% real interest rate.

FIGURE II: WTP and Cost Components for Medicaid Expansions to Pregnant Women and Infants



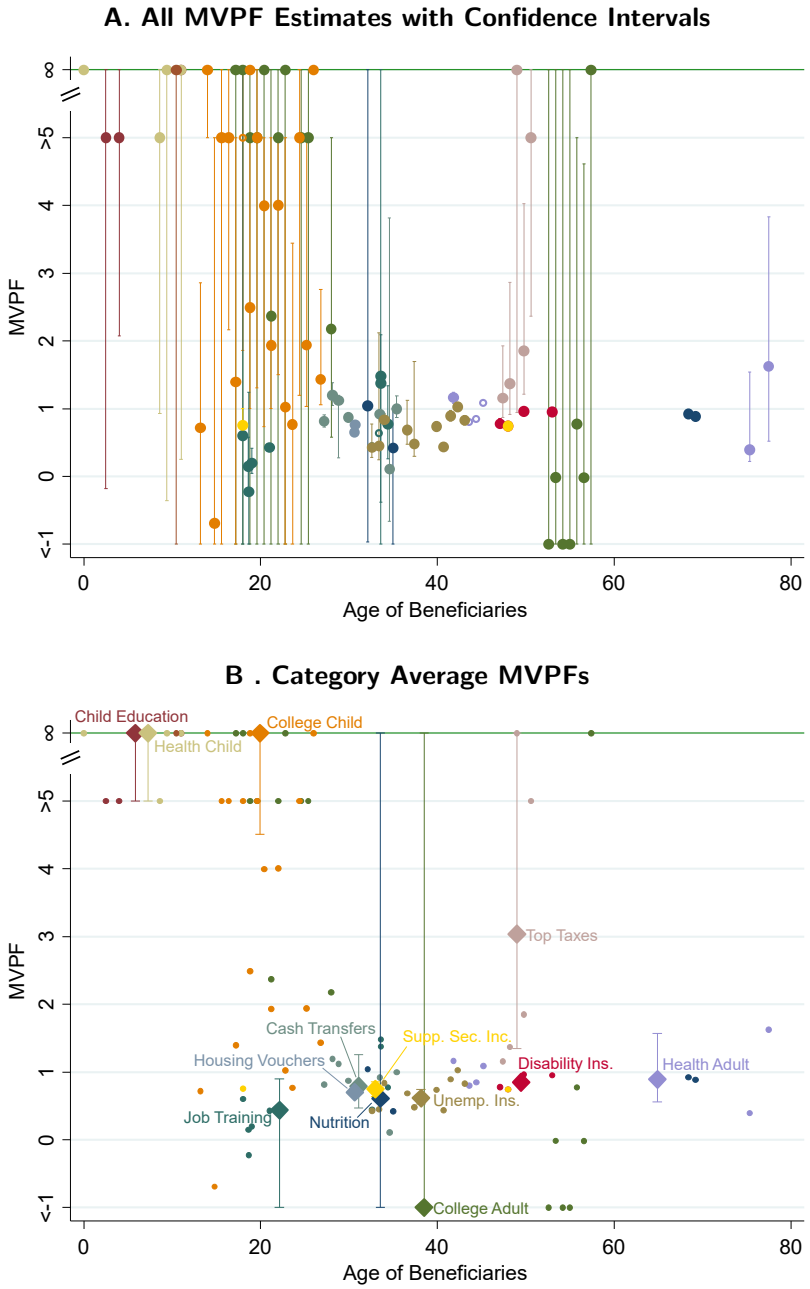
Notes: This figure illustrates the cost to the government of providing Medicaid to pregnant women and infants. The evidence comes from State Medicaid expansions between 1979 and 1992. Panel A breaks the total cost down into its various components. The savings on uncompensated care come from Currie and Gruber (1996), who estimate rates of uninsurance, and Gold and Kenney (1985) who estimate the quantity of uncompensated care for the uninsured. The savings on future health costs come from Miller and Wherry (2018). The increase in government revenue combines an effective tax rate with the estimates of earnings gains from Miller and Wherry (2018). Panel B reports the components of our WTP calculations. The conservative WTP is given by the reduction in spending on other health insurance payments. The point estimate includes the willingness to pay for reductions in infant mortality, combined with the change in income for children over their life cycle after taxes and educational expenses. All numbers in 2011 dollars deflating using the CPI-U-RS and discounted using a 3% real interest rate.

FIGURE III: MVPF Estimates by Age of Policy Beneficiary



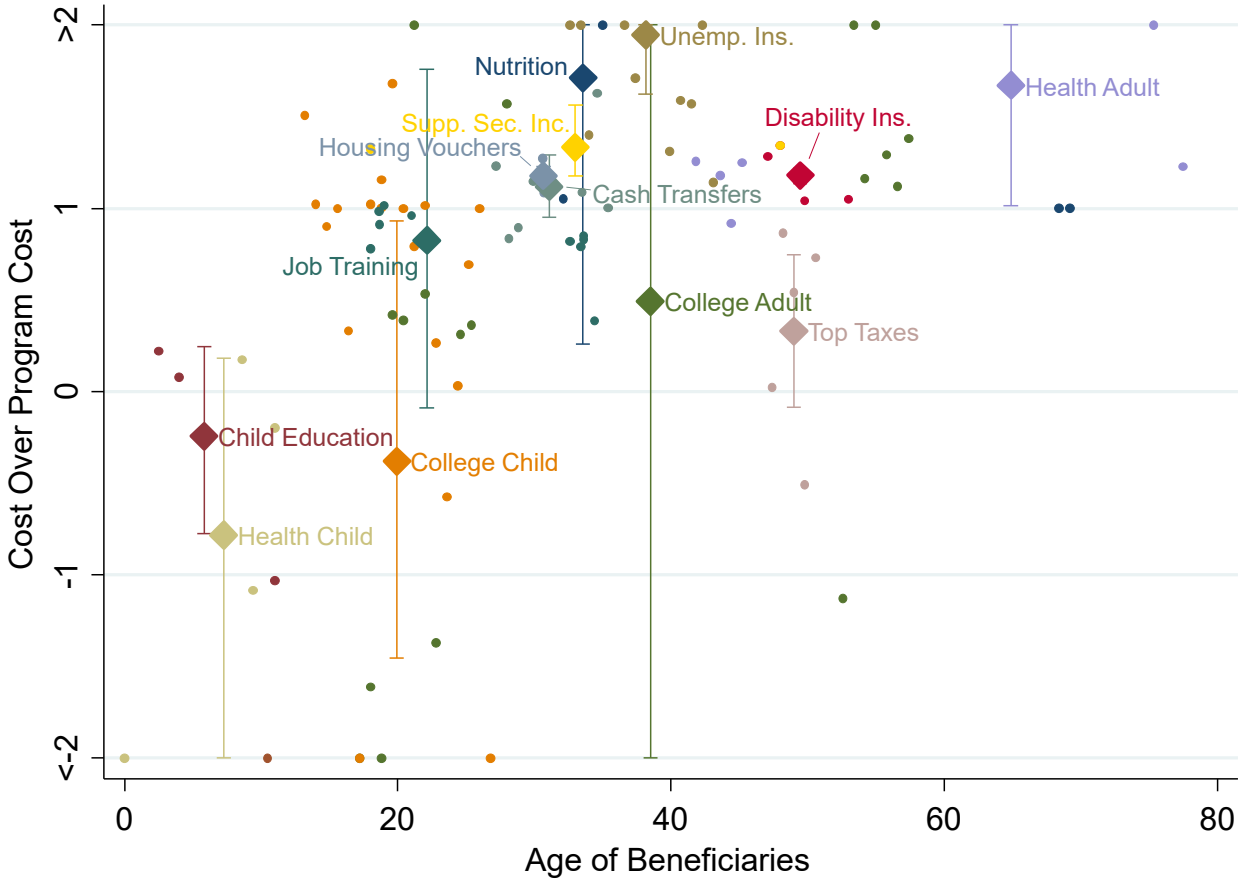
Notes: This figure presents MVPF estimates for all policies in our baseline sample. For each MVPF, we plot them as a function of the average age of the policy's beneficiaries. In cases where both parents and children potentially benefit, we assign the age of the individuals with the highest willingness to pay. Where policies within a category have the same age, we stagger these ages around this common value for clarity. On the vertical axis, we report the MVPF estimates, capping these estimates at 5. We separately report cases where the MVPF is infinite on the uppermost line in green.

FIGURE IV: MVPF Estimates and Category Averages with Confidence Intervals



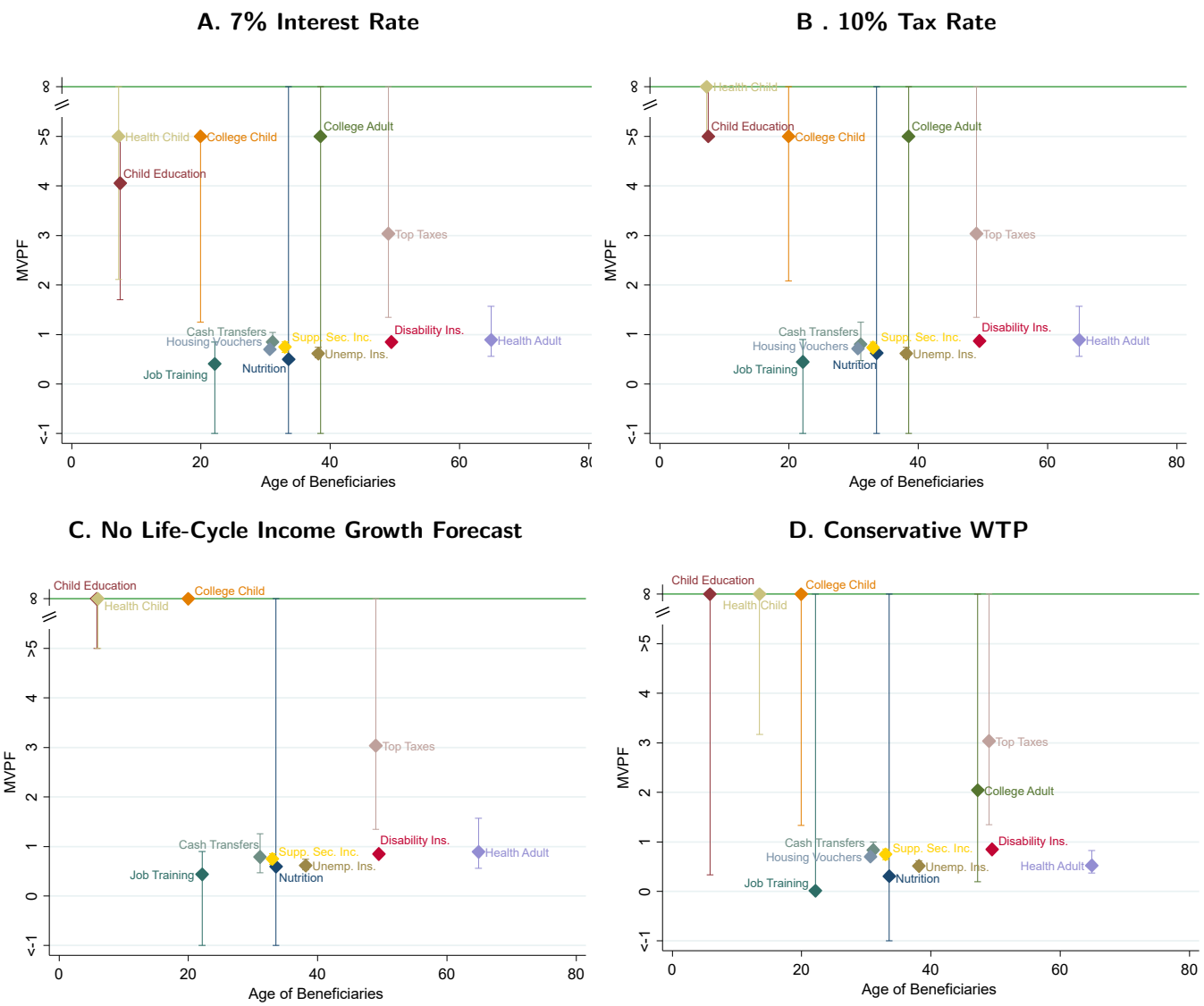
Notes: Panel A presents the MVPFs and 95% confidence intervals for each policy in our baseline sample, plotted as a function of the average age of the policy's beneficiaries. Panel B presents $1 spend domain averages and 95% confidence intervals across categories of programs, plotted as a function of the average age of each policy's beneficiaries within category. Individual policy MVPFs are shown in smaller dots, color-coded to align with their respective categories. In both panels, we report the MVPF estimates on the vertical axis, capping these estimates at 5 and separately reporting cases where the MVPF is infinite on the uppermost line in green. All confidence intervals are 95% bootstrapped confidence intervals with adjustments discussed in Appendix H.

FIGURE V: Net Government Costs per Dollar of Programmatic Spending



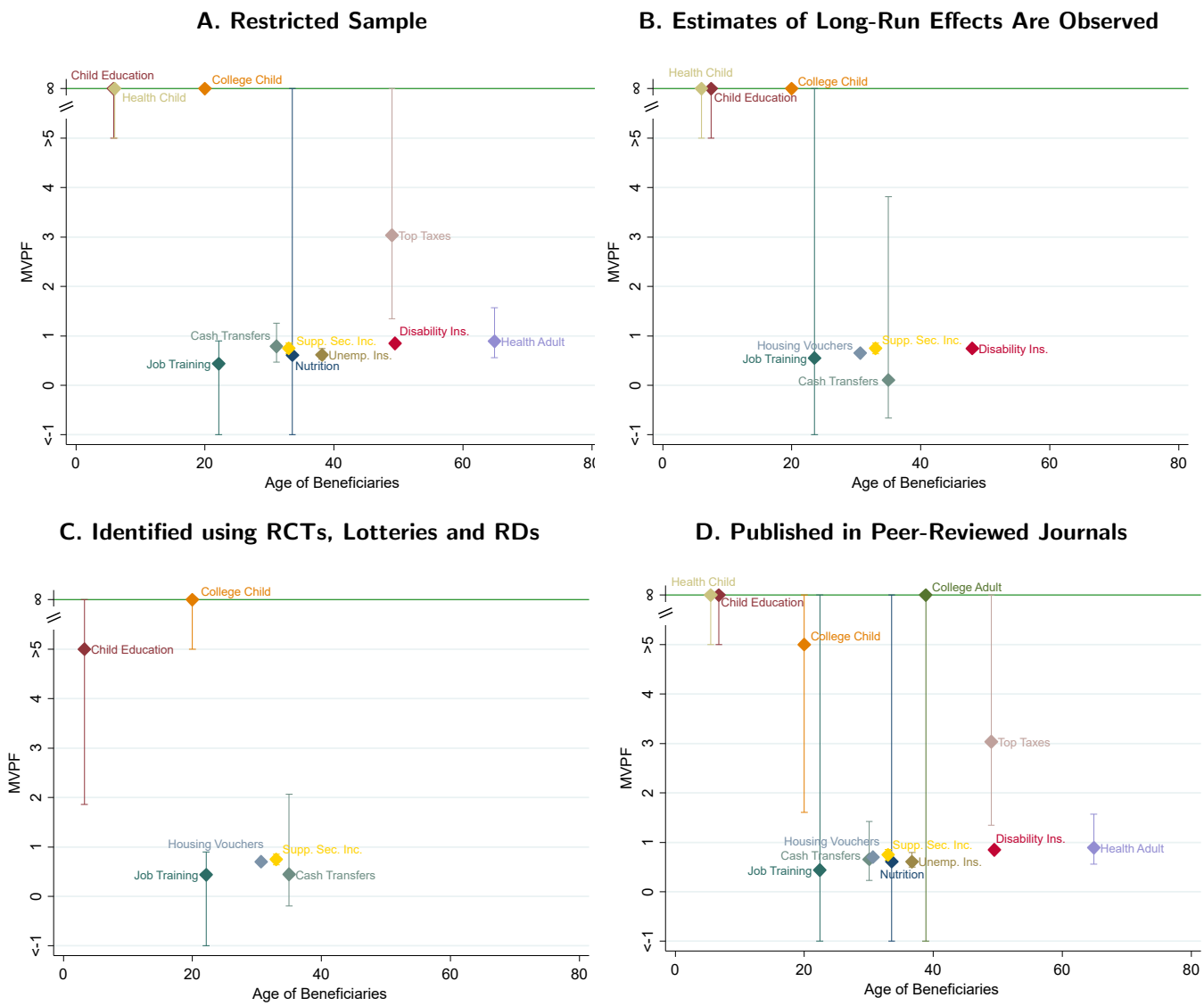
Notes: This figure presents estimates of costs normalized by initial programmatic for each category-average group of policies in our baseline sample. We plot these estimates as a function of the average age of each policy's beneficiaries within category. Bootstrapped 95% confidence intervals with adjustments discussed in Appendix H are shown for the category averages. The normalized costs of individual policies are shown in smaller dots, color-coded to align with their respective categories.

FIGURE VI: Specification Robustness



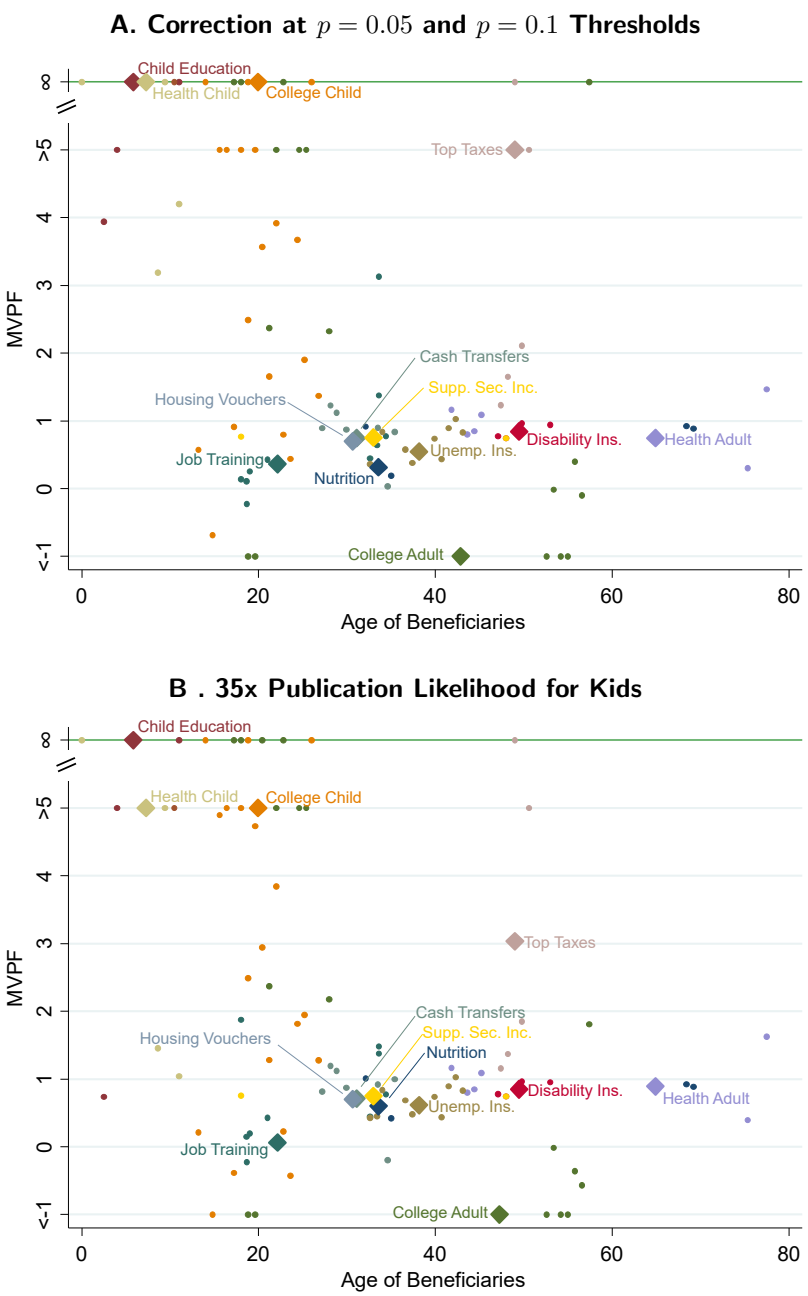
Notes: This figure presents the category-average MVPFs from Figure III using a range of different alternative specifications that are more conservative than our baseline specifications. Panel A replaces our assumption of a 3% real interest rate with a 7% real interest rate. Panel B replaces our assumption of a CBO tax rate (often around 18-20%) with a 10% tax rate. Panel C replaces our baseline income projection procedure with a procedure that assumes zero income growth over the lifecycle. We use our restricted sample of policies for this specification. See appendix I for further details. Panel D replaces our point estimate WTP measures with our conservative measures of WTP. We report bootstrapped 95% confidence intervals with adjustments discussed in Appendix H for each category average.

FIGURE VII: Sample Restrictions



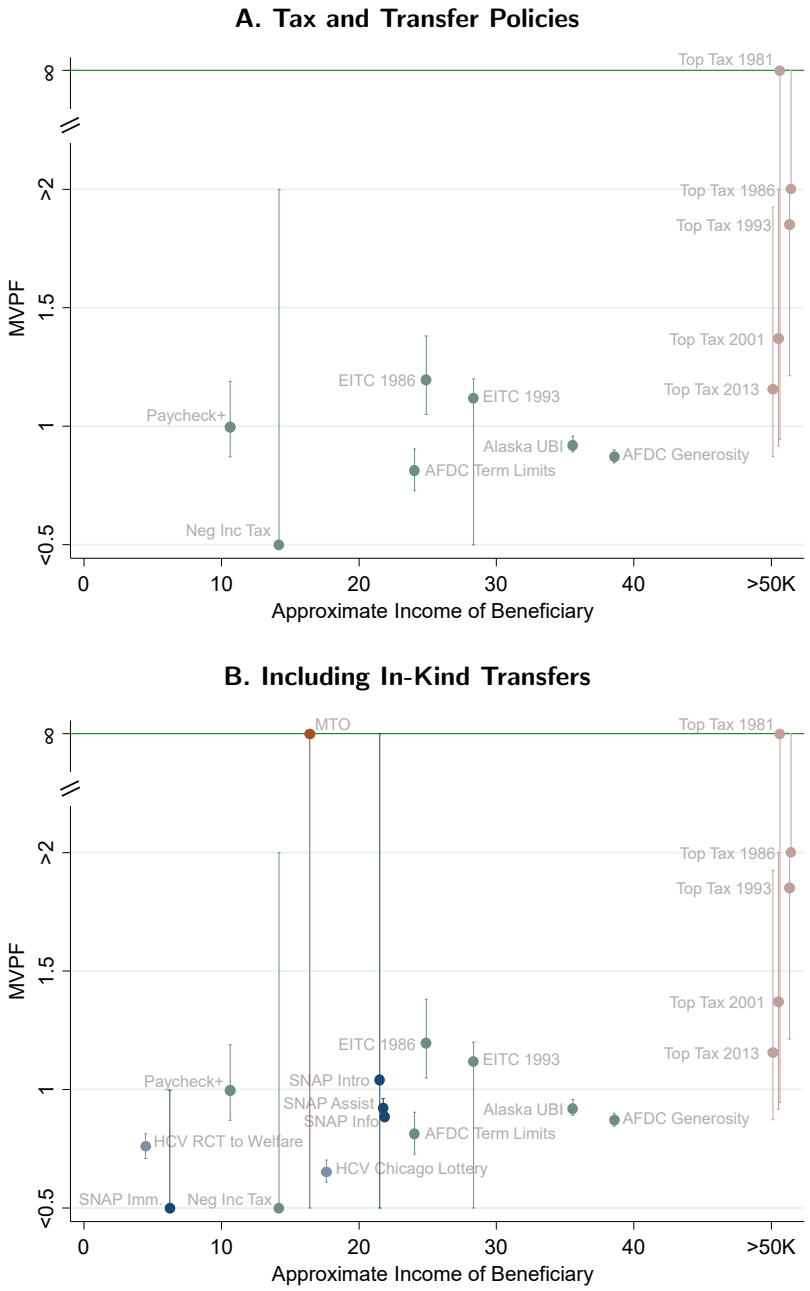
Notes: This figure presents the category-average MVPFs from Figure III using a range of alternative sample restrictions. Panel A considers our restricted sample that drops estimates for which we are forecasting earnings impacts based on a policy's impact on college attendance. Panel B restricts the sample to only policies for which earnings outcomes are estimated for at least 5 years of follow-up after the policy. For this panel we show group averages even for groups with a single policy. Panel C restricts the sample to policies whose identification strategy is an randomized control trial, lottery, or regression discontinuity. Panel D restricts to policies whose primary analyses have been published in a peer-reviewed journal. We report bootstrapped 95% confidence intervals with adjustments discussed in Appendix H for each category average.

FIGURE VIII: Publication Bias



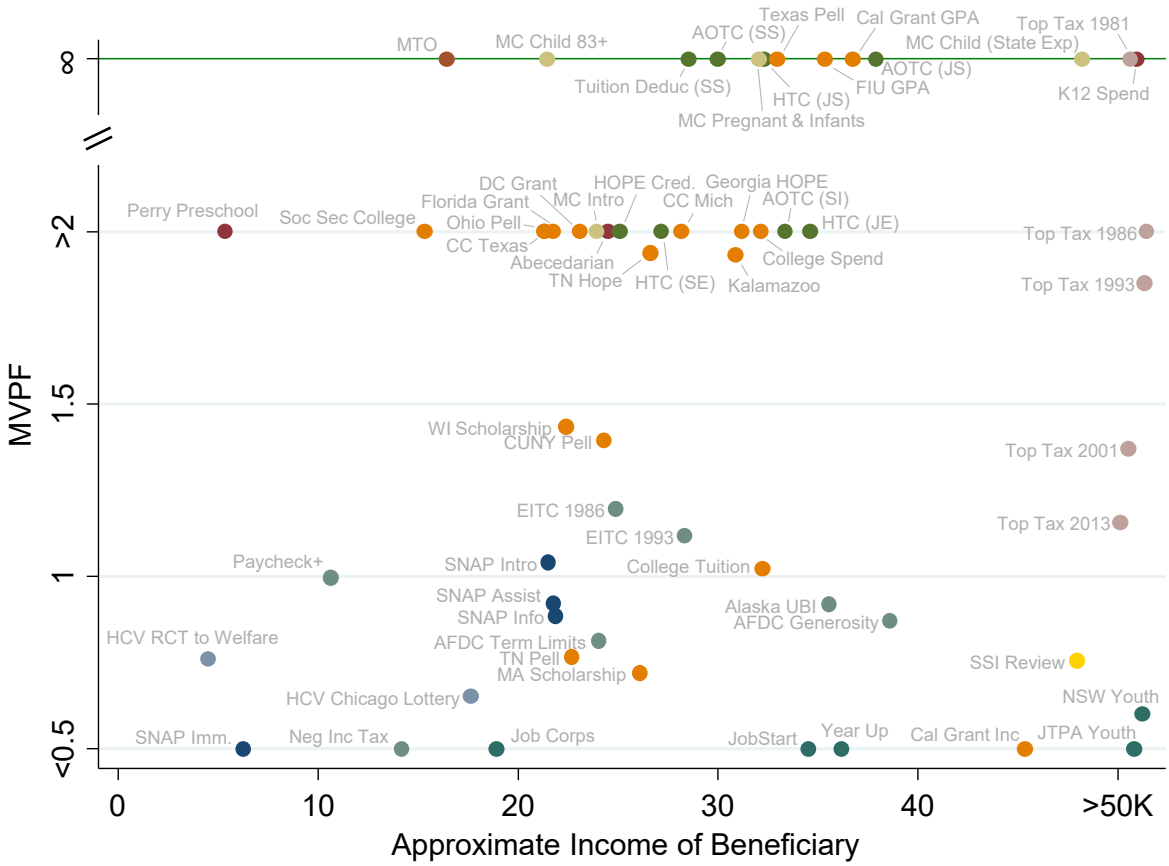
Notes: This figure presents the MVPF estimates from Figure III and category averages in Figure IVB using estimates corrected for publication bias from the method of Andrews and Kasy (2018). Panel A reports estimates using the corrections using the publication likelihood estimated from our model that imposes jumps at $p = 0.05$ and $p = 0.10$, as shown in Table III columns (3) and (6). In panel B we report corrected estimates under an assumption that child policies are 35 time more likely to be published if they find a positive effect on children's outcomes (and we assume no publication bias for adult policies or for child policies that find negative effects on children). This 35 times corresponds to the estimated publication bias implied by a large-scale replication of experimental economics papers by Camerer et al. (2016) (Table 1 of Andrews and Kasy (2018) reports that insignificant results are 0.029 times as likely to be published). We do not report confidence intervals for these estimates (to our knowledge there is no well-accepted method of constructing such intervals); but we refer readers to Figure IVB to note that some of these category averages are imprecise.

FIGURE IX: MVPF by Income of Beneficiaries



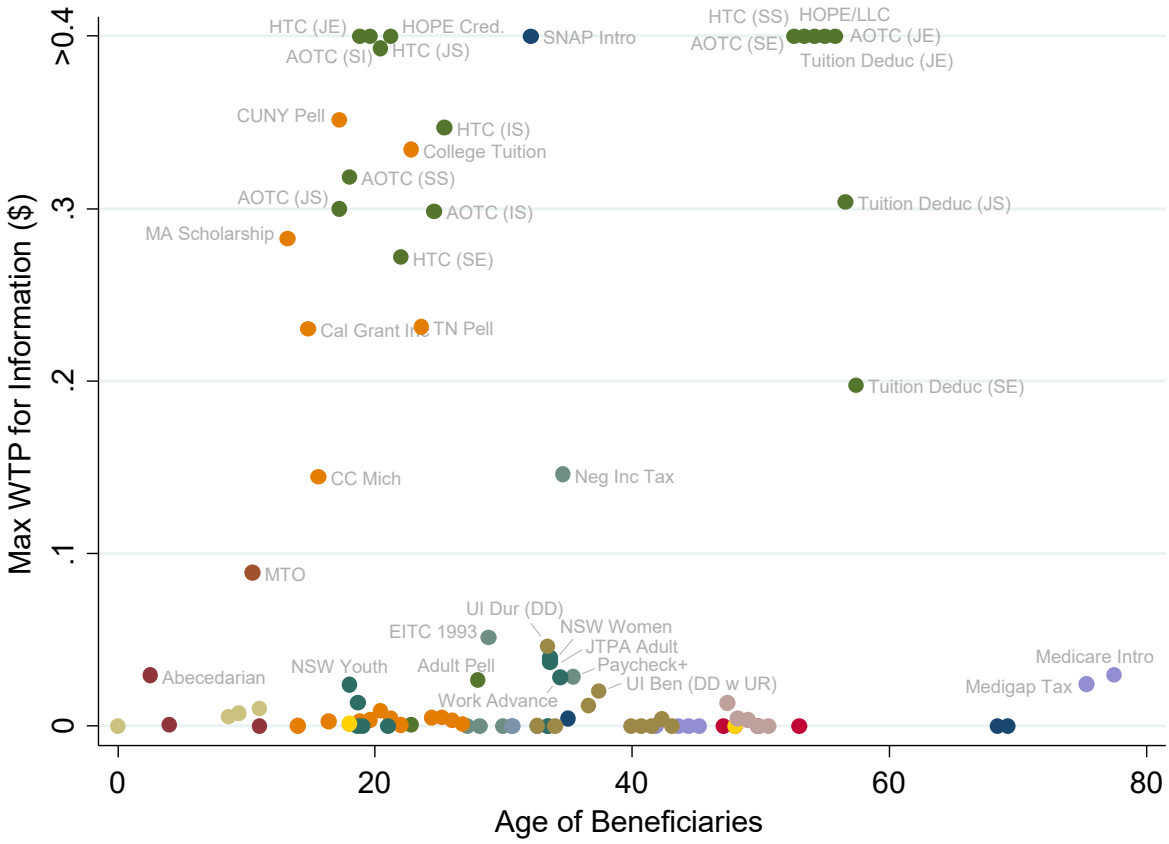
Notes: Panel A shows MVPFs for tax and transfer policies in our baseline sample against the income of their economic beneficiaries. Panel B adds in-kind transfer policies. All confidence intervals are 95% bootstrapped confidence intervals with adjustments discussed in Appendix H. See Figure III for an explanation of the color scheme.

FIGURE X: MVPF by Income of Beneficiaries: Including Child Policies



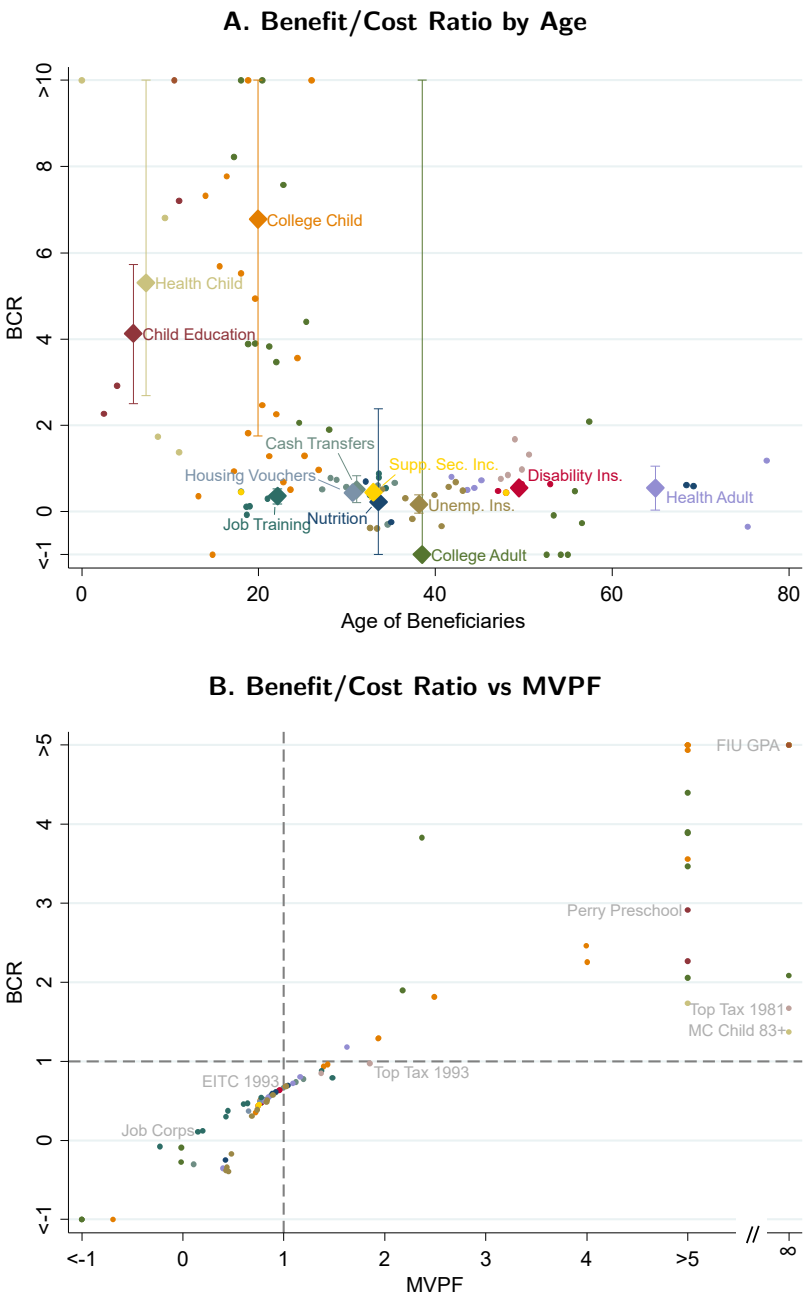
Notes: This Figure MVPFs as a function of the average income of beneficiaries for tax and transfer policies (shown in Figure IXA) combined with our estimates for in-kind transfer policies and child policies (Child education, college, health, and job training). The income measures should be considered approximations, as not all papers report consistent measures of incomes of their samples. We include all papers for which we are able to obtain a measure of income of the beneficiaries, and we attempt to normalize each of these measures to correspond to a notion of individual income per adult in the household at age 30. All confidence intervals are 95% bootstrapped confidence intervals with adjustments discussed in Appendix H. See Figure III for an explanation of the color scheme.

FIGURE XI: Value of Information



Notes: This figure presents the value of information, v^{info}, discussed in Section VI.A, for each policy in our sample as a function of the average age of the policy beneficiaries. See Figure III for an explanation of the color scheme.

FIGURE XII: Comparison to CBA



Notes: This figure presents estimates of benefit-cost ratios for all policies evaluated in the paper and shows their relationship to the MVPF. The method for calculating these benefit-cost ratios is outlined in Section 2. We assume a marginal deadweight loss of $\phi = 50\%$ for these calculations. Panel A plots the benefit-cost ratio of each policy as a function of the age of the beneficiaries, along with category average estimates and their confidence intervals. The capped lines show the 95% bootstrapped confidence intervals with adjustments discussed in Appendix H. Panel B plots the benefit-cost ratio of each policy as a function of the MVPF estimate for the policy. See Figure III for an explanation of the color scheme.

Table 1: Details of All Programs Studied

Program	Label	Year Implemented	Age of Beneficiaries	Sample: Baseline	Sample: Restricted	Sample: Extended	Paper Estimates MVPF	Papers Utilized
Panel A: Education and Job Training								
Child Education								
Carolina Abecedarian Study	Abecedarian	1975	3	x	x	x		Barnett and Masse (2007) Campbell et al. (2012) Helburn (1995) Masse and Barnett (2002) Masse (2003)
Chicago Child-Parent Centers, Extended Program	CPC Extended	1985	6			x		Reynolds et al. (2002) Reynolds et al. (2011)
Chicago Child-Parent Centers, Preschool Program	CPC Preschool	1983	4			x		Reynolds et al. (2002) Reynolds et al. (2011)
Chicago Child-Parent Centers, School Age Program	CPC School	1986	8			x		Reynolds et al. (2002) Reynolds et al. (2011)
Head Start Impact Based on Geographic Availability	Head Start Geo	1970	4			x		Ludwig and Miller (2007)
Head Start Impact Study	Head Start RCT	2002	3			x	x	Walters and Kline (2016)
K-12 School Finance Reform	K12 Spend	1991	11	x	x	x		Jackson et al. (2016)
K-12 School Spending in Michigan	K12 Spend Mich.	1994	11			x		Hyman (2017)
Perry Preschool Program	Perry Preschool	1962	4	x	x	x		Heckman et al. (2010) Heckman et al. (2011)
College Adult								
American Opportunity Tax Credit, Independent Single Filers at Phase Start	AOTC (IS)	2011	25	x		x		Bulman and Hoxby (2015)
American Opportunity Tax Credit, Joint Filers at Phase End	AOTC (JE)	2011	55	x		x		Bulman and Hoxby (2015)
American Opportunity Tax Credit, Joint Filers at Phase Start	AOTC (JS)	2011	20	x		x		Bulman and Hoxby (2015)
American Opportunity Tax Credit, Simulated Instrument	AOTC (SI)	2009	20	x		x		Bulman and Hoxby (2015)
American Opportunity Tax Credit, Single Filers at Phase End	AOTC (SE)	2011	55	x		x		Bulman and Hoxby (2015)
American Opportunity Tax Credit, Single Filers at Phase Start	AOTC (SS)	2011	20	x		x		Bulman and Hoxby (2015)
Hope Tax Credit	HOPE Cred.	1999	20	x		x		Turner (2011)
Hope Tax Credit, Indpendent Single Filers at Phase Start	HTC (IS)	2007	25	x		x		Bulman and Hoxby (2015)
Hope Tax Credit, Joint Filers at Phase End	HTC (JE)	2007	20	x		x		Bulman and Hoxby (2015)
Hope Tax Credit, Joint Filers at Phase Start	HTC (JS)	2007	20	x		x		Bulman and Hoxby (2015)
Hope Tax Credit, Single Filers at Phase End	HTC (SE)	2007	20	x		x		Bulman and Hoxby (2015)
Hope Tax Credit, Single Filers at Phase Start	HTC (SS)	2007	55	x		x		Bulman and Hoxby (2015)
Hope and Lifetime Learners Tax Credits	HOPE/LLC	1998	55	x		x		Long (2004)
Pell Grants Introduction to Adults	Adult Pell	1973	28	x		x		Seftor and Turner (2002)
Tax Deduction for Postsecondary Tuition, Joint Filers at Phase End	Tuition Deduc (JE)	2006	55	x		x		Hoxby and Bulman (2016) LaLumia (2012)
Tax Deduction for Postsecondary Tuition, Joint Filers at Phase Start	Tuition Deduc (JS)	2006	55	x		x		Hoxby and Bulman (2016) LaLumia (2012)
Tax Deduction for Postsecondary Tuition, Single Filers at Phase End	Tuition Deduc (SE)	2006	55	x		x		Hoxby and Bulman (2016) LaLumia (2012)
Tax Deduction for Postsecondary Tuition, Single Filers at Phase Start	Tuition Deduc (SS)	2006	20	x		x		Hoxby and Bulman (2016) LaLumia (2012)
College Child								
Cal Grant, GPA Threshold	Cal Grant GPA	1998	20	x	x	x		Bettinger et al. (2019)
Cal Grant, Income Threshold	Cal Grant Inc	1998	20	x	x	x		Bettinger et al. (2019)
City University of New York Pell Grants	CUNY Pell	2009	20	x		x		Marx and Turner (2018)
Community College Tuition Changes, Michigan	CC Mich	2005	20	x		x		Acton (2018)
Community College Tuition Changes, Texas	CC Texas	2005	20	x		x		Denning (2017)
District of Columbia Tuition Assistance Grant Program	DC Grant	1999	20	x		x		Abraham and Clark (2006)
Florida International University Admissions at GPA Threshold	FIU GPA	1999	20	x	x	x		Zimmerman (2014)
Florida Student Access Grant	Florida Grant	2001	20	x		x		Castleman and Long (2016)
Free Application for Federal Student Aid, Dependent Year Impact	Free FAFSA (Dep)	2008	20			x		Bettinger et al. (2012)

							U.S. Department of Education Office of Postsecondary Education (2010)
Free Application for Federal Student Aid, Independent Year Impact	Free FAFSA (Indep)	2008	20			x	Bettinger et al. (2012)
							U.S. Department of Education Office of Postsecondary Education (2010)
Georgia HOPE Scholarship	Georgia HOPE	1995	20	x		x	Cornwell et al. (2006)
							Cornwell et al. (2003)
Hail Michigan Aid Awareness Letter	HAIL Aid	2016	20			x	Dynarski et al. (2018)
							Hoekstra (2009)
Kalamazoo Promise Scholarship	Kalamazoo	2006	20	x		x	Bartik et al. (2017)
							Bartik et al. (2016)
Massachussetts Adams Scholarship	MA Scholarship	2005	20	x		x	Cohodes and Goodman (2014)
							Goodman (2008)
Pell Grants in Ohio	Ohio Pell	2000	19	x		x	Bettinger (2004)
Pell Grants in Tennessee	TN Pell	2008	20	x		x	Carruthers and Welch (2019)
							U.S. Department of Education Office of Postsecondary Education (2010)
Pell Grants in Texas	Texas Pell	2008	20	x	x	x	Denning et al. (2017)
Social Security Student Benefit Program	Soc Sec College	1982	20	x		x	Dynarski (2003)
Spending at Colleges from State Appropriations	College Spend	2001	20	x		x	Deming and Walters (2017)
Tennessee HOPE Scholarships	TN Hope	2008	20	x		x	Bruce and Carruthers (2014)
Tuition Cuts at Colleges from State Appropriations	College Tuition	2001	20	x		x	Deming and Walters (2017)
Wisconsin Scholar Grant to Low-Income College Students	WI Scholarship	2009	20	x		x	Goldrick-Rab et al. (2016)
Job Training							
Job Corps	Job Corps	1995	19	x	x	x	Schochet (2018)
							Schochet et al. (2008)
							Schochet et al. (2006)
Job Training Partnership Act, Adults	JTPA Adult	1988	34	x	x	x	Bloom et al. (1997)
Job Training Partnership Act, Youth	JTPA Youth	1988	19	x	x	x	Bloom et al. (1997)
JobStart	JobStart	1986	19	x	x	x	Cave et al. (1993)
National Supported Work Demonstration, Adult Women	NSW Women	1976	34	x	x	x	Couch (1992)
							Hollister et al. (1984)
National Supported Work Demonstration, Ex-Addicts	NSW Ex-Addict	1976	33	x	x	x	Hollister et al. (1984)
National Supported Work Demonstration, Ex-Offenders	NSW Ex-Offender	1976	33	x	x	x	Hollister et al. (1984)
National Supported Work Demonstration, Youth	NSW Youth	1976	18	x	x	x	Couch (1992)
							Hollister et al. (1984)
Work Advance	Work Advance	2012	34	x	x	x	Hendra et al. (2016)
							Schaberg (2017)
Year Up	Year Up	2013	21	x	x	x	Fein and Hamadyk (2018)
Panel B: Social Insurance							
Disability Insurance							
Disability Insurance Changes in Benefit Generosity	DI Generosity	2004	50	x	x	x	Gelber et al. (2017)
Disability Insurance Judge Leniency	DI Judge	2005	47	x	x	x	Gelber et al. (2017)
							Maestas et al. (2013)
Disability Insurance Medical Examiner Leniency	DI Examiner	1995	48	x	x	x	French and Song (2014)
							Gelber et al. (2017)
Disability Insurance to Veterans	DI Veterans	2001	53	x	x	x	Autor et al. (2016)
Health Adult							
Health Insurance Subsidies in Massachusetts to Individuals at 150% of the Federal Poverty Line	Mass HI (150%FPL)	2011	44	x	x	x	Finkelstein et al. (2017)
							Hendren (2017c)
Health Insurance Subsidies in Massachusetts to Individuals at 200% of the Federal Poverty Line	Mass HI (200%FPL)	2011	44	x	x	x	Finkelstein et al. (2017)
							Hendren (2017c)
Health Insurance Subsidies in Massachusetts to Individuals at 250% of the Federal Poverty Line	Mass HI (250%FPL)	2011	44	x	x	x	Finkelstein et al. (2017)
							Hendren (2017c)
Medicare Introduction in 1965	Medicare Intro	1965	78	x	x	x	United States Census Bureau (1966)
							Finkelstein and McKnight (2008)
Oregon Health Insurance Experiment (Provided to Single Adults)	Oregon Health	2008	42	x	x	x	Finkelstein et al. (2012)
							Finkelstein et al. (Forthcoming)

Taxation of Medigap Policies	Medigap Tax	2002	75	x	x	x		Cabral and Mahoney (2019)
Health Child								
Medicaid Expansion to Children Born after September 30, 1983	MC Child 83+	1990	11	x		x	x	Lo Sasso and Seamster (2007) Meyer and Wherry (2016) Wherry et al. (2018)
Medicaid Expansions to Pregnant Women & Infants	MC Pregnant & Infants	1986	0	x	x	x		Dave et al. (2015) Currie and Gruber (1996) Miller and Wherry (2018)
Medicaid Expansions to Young Children	MC Child (State Exp)	1986	9	x	x	x		Boudreaux et al. (2016) Brown et al. (2017)
Medicaid Introduction to AFDC-eligible Families	MC Intro	1968	9	x	x	x	x	Goodman-Bacon (2017)
Supplemental Security Income								
Supplemental Security Income Age 18 Medical Review	SSI Review	1996	18	x	x	x	x	Deshpande (2016)
Supplemental Security Income Medical Examiner Leniency	SSI Judge	2005	48	x	x	x		Deshpande (2016) French and Song (2014) Gelber et al. (2017) United States Social Security Administration (2014)
Unemployment Insurance								
Unemployment Insurance Benefit Changes (Diff in Diff Across States in Chetty (2008))	UI Ben (State Max)	1992	37	x	x	x		Chetty (2008) Gruber (1997) Hendren (2017b) Schmieder and Von Wachter (2016)
Unemployment Insurance Benefit Changes (Diff in Diff Across States in Katz and Meyer (1990))	UI Ben (DD)	1980	33	x	x	x		Gruber (1997) Hendren (2017b) Katz and Meyer (1990) Schmieder and Von Wachter (2016)
Unemployment Insurance Benefit Changes (Diff in Diff Across States in Kroft and Notowodigdo (2016))	UI Ben (DD w UR)	1992	37	x	x	x		Gruber (1997) Hendren (2017b) Kroft and Notowidigdo (2016) Schmieder and Von Wachter (2016)
Unemployment Insurance Benefit Changes in Georgia	UI Ben (GA)	1979	42	x	x	x		Gruber (1997) Hendren (2017b) Schmieder and Von Wachter (2016) Solon (1985)
Unemployment Insurance Benefit Changes in Missouri (Expansion Estimates)	UI Ben (MO Exp)	2005	42	x	x	x		Card et al. (2015) Gruber (1997) Hendren (2017b) Schmieder and Von Wachter (2016)
Unemployment Insurance Benefit Changes in Missouri (Recession Estimates)	UI Ben (MO Rec)	2010	42	x	x	x		Card et al. (2015) Gruber (1997) Hendren (2017b) Schmieder and Von Wachter (2016)
Unemployment Insurance Benefit Changes in New York	UI Ben (NY)	1989	42	x	x	x		Gruber (1997) Hendren (2017b) Meyer and Mok (2007) Schmieder and Von Wachter (2016)
Unemployment Insurance Benefit Changes via Regression Kink in Benefit Schedule	UI Ben (RK)	1980	34	x	x	x		Gruber (1997) Hendren (2017b) Landais (2015) Schmieder and Von Wachter (2016)
Unemployment Insurance Duration Extensions (Diff in Diff Across States in Katz and Meyer (1990))	UI Dur (DD)	1980	33	x	x	x		Ganong and Noel (2019) Gruber (1997) Hendren (2017b) Katz and Meyer (1990) Schmieder and Von Wachter (2016)
Unemployment Insurance Duration Extensions in Missouri	UI Dur (MO)	2011	42	x	x	x		Ganong and Noel (2019) Gruber (1997)

								Hendren (2017b) Johnston and Mas (2018) Schmieder and Von Wachter (2016)
Panel C: In-Kind Transfers								
Housing Vouchers								
Effects of Housing Vouchers on AFDC Families Experiment	HCV RCT to Welfare	2000	31	x		x		Jacob and Ludwig (2012) Mills et al. (2006) Wood et al. (2008)
Housing Vouchers in Chicago	HCV Chicago Lottery	1997	31	x		x		Jacob and Ludwig (2012) Jacob et al. (2014)
Jobs Plus	Jobs+	1998	35			x		Bloom et al. (2005) Riccio (2006)
MTO								
Moving to Opportunity Experiment Providing Vouchers and Counseling	MTO	1996	10	x	x	x		Chetty et al. (2016) Goering et al. (1999) Sanbonmatsu et al. (2011)
Nutrition								
Special Supplemental Nutrition Program for Women, Infants, and Children	WIC	1975	26			x		Black et al. (2007) Hoynes et al. (2011) Whitmore (2002)
Supplemental Nutrition Assistance Program Application Assitance	SNAP Assist	2016	69	x	x	x	x	Finkelstein and Notowidigdo (2019)
Supplemental Nutrition Assistance Program Application Information	SNAP Info	2016	69	x	x	x	x	Finkelstein and Notowidigdo (2019)
Supplemental Nutrition Assistance Program Introduction	SNAP Intro	1968	32	x	x	x		Hoynes et al. (2016) Almond et al. (2011) Bailey et al. (2019) Hoynes et al. (2011) Hoynes and Schanzenbach (2012)
Supplemental Nutrition Assistance Program for Immigrants	SNAP Imm.	2001	35	x	x	x		East (2018) Finkelstein and Notowidigdo (2019)
Panel D: Taxes and Cash Transfers								
Cash Transfers								
1986 Earned Income Tax Credit Expansion	EITC 1986	1986	28	x	x	x		Ackerman et al. (2009) Blank and Ruggles (1996) Hotz and Scholz (2003) Meyer and Rosenbaum (2001) Moffitt (2002) Scholz (1993) Crouse and Waters (2014) Eissa and Hoynes (2004) Eissa and Liebman (1996) Tax Policy Center (2016)
1993 Earned Income Tax Credit Expansion	EITC 1993	1993	29	x	x	x		Dahl and Lochner (2012) Meyer and Rosenbaum (2001) Bastian and Michelmore (2018) Chetty et al. (2011) Crouse and Waters (2014) Eissa and Hoynes (2004) Hoynes and Patel (2018) Manoli and Turner (2018) Maxfield (2013) Michelmore (2013) Tax Policy Center (2016)
Aid to Families with Dependent Children (Greater Benefit Generosity)	AFDC Generosity	1990	30	x	x	x		Ackerman et al. (2009) Hotz and Scholz (2003) Meyer and Rosenbaum (2001) Currie and Cole (1993)

							Moffitt (2003)
Aid to Families with Dependent Children (Term Limit Modifications)	AFDC Term Limits	1996	27	x	x	x	Grogger (2003)
							Pavetti (1995)
Alaska Permanent Fund Dividend	Alaska UBI	1998	34	x	x	x	Ackerman et al. (2009)
							Bhargava and Manoli (2015)
							Blank and Ruggles (1996)
							Hotz and Scholz (2003)
							Jones and Marinescu (2018)
							Meyer and Rosenbaum (2001)
							Moffitt (2002)
Paycheck Plus Experiment Providing EITC-benefits to Adults without Dependents	Paycheck+	2013	35	x	x	x	Miller et al. (2017)
Seattle-Denver Income Maintenance Experiment	Neg Inc Tax	1971	35	x	x	x	Price and Song (2016)
							Administration (2018)
							Tax Foundation (2013)
							U.S. Department of Health & Human Service
							Von Wachter et al. (2011)
Top Taxes							
Top Tax 2013 Increases from Affordable Care Act	Top Tax 2013	2013	49	x	x	x	Kawano et al. (2016)
							Hendren (2017a)
Top Tax Rate Increase in Omnibus Budget Reconciliation Act 1993	Top Tax 1993	1993	49	x	x	x	Atkinson et al. (2011)
							Carroll (1998)
Top Tax Rate Reductions in Tax Reform Act of 1986	Top Tax 1986	1986	49	x	x	x	Atkinson et al. (2011)
							Auten and Carroll (1999)
Top Taxes, Economic Growth And Tax Relief Reconciliation Act 2001	Top Tax 2001	2001	49	x	x	x	Atkinson et al. (2011)
							Heim (2009)
Top Taxes, Economic Recovery Tax Act 1981	Top Tax 1981	1981	49	x	x	x	Atkinson et al. (2011)
							Saez (2003)
Panel E: Welfare Reform							
Welfare Reform							
Welfare to Work Alameda Mandatory Mixed-Initial-Activity Programs	GAIN Alm.	1988	31			x	Freedman et al. (1996)
							Greenberg et al. (2010)
Welfare to Work Atlanta Mandatory Education-First Programs	HCD NEWWS Atl.	1992	33			x	Greenberg et al. (2010)
							Hamilton et al. (2001)
Welfare to Work Atlanta Mandatory Job-Search-First Programs	LFA NEWWS Atl.	1992	33			x	Greenberg et al. (2010)
							Hamilton et al. (2001)
Welfare to Work Butte Mandatory Mixed-Initial-Activity Programs	GAIN Butte	1987	31			x	Freedman et al. (1996)
							Greenberg et al. (2010)
Welfare to Work Columbus Integrated Mandatory Education-First Programs	NEWWS Col. Int.	1992	32			x	Greenberg et al. (2010)
							Hamilton et al. (2001)
Welfare to Work Columbus Traditional Mandatory Education-First Programs	NEWWS Col. Trad.	1992	32			x	Greenberg et al. (2010)
							Hamilton et al. (2001)
Welfare to Work Connecticut Time-Limit-Mix Programs	Jobs First	1996	31			x	Bloom et al. (2002)
							Greenberg et al. (2010)
Welfare to Work Cook County Mandatory Work Experience Program	WIN Demo.	1985	32			x	Brock et al. (1993)
							Greenberg et al. (2010)
Welfare to Work Detroit Mandatory Education-First Programs	NEWWS Det.	1992	30			x	Greenberg et al. (2010)
							Hamilton et al. (2001)
Welfare to Work Florida Mandatory Mixed-Initial-Activity Programs	Proj. Ind. FL	1990	32			x	Greenberg et al. (2010)
							Kemple et al. (1995)
Welfare to Work Florida Time-Limit-Mix Programs	FTP	1994	29			x	Bloom et al. (2000)
							Greenberg et al. (2010)
Welfare to Work Grand Rapids Mandatory Education-First Programs	HCD NEWWS Gr. Rap.	1991	28			x	Greenberg et al. (2010)
							Hamilton et al. (2001)
Welfare to Work Grand Rapids Mandatory Job-Search-First Programs	LFA NEWWS Gr. Rap.	1991	28			x	Greenberg et al. (2010)
							Hamilton et al. (2001)
Welfare to Work Los Angeles Mandatory Job-Search-First Programs	GAIN LA jobs	1996	34			x	Freedman et al. (2000)
							Greenberg et al. (2010)
Welfare to Work Los Angeles Mandatory Mixed-Initial-Activity Programs	GAIN LA	1988	31			x	Freedman et al. (1996)
							Greenberg et al. (2010)
Welfare to Work Minnesota Earnings Supplement Programs	MFIP	1994	29			x	Greenberg et al. (2010)

					Miller et al. (2000)
Welfare to Work Portland Mandatory Mixed-Initial-Activity Programs	NEWWS Port.	1993	30	x	Greenberg et al. (2010)
					Hamilton et al. (2001)
Welfare to Work Riverside Mandatory Education-First Programs	HCD NEWWS Riv.	1991	32	x	Greenberg et al. (2010)
					Hamilton et al. (2001)
Welfare to Work Riverside Mandatory Job-Search-First Programs	LFA NEWWS Riv.	1991	32	x	Greenberg et al. (2010)
					Hamilton et al. (2001)
Welfare to Work Riverside Mandatory Mixed-Initial-Activity Programs	GAIN Riv.	1987	31	x	Freedman et al. (1996)
					Greenberg et al. (2010)
Welfare to Work San Diego Mandatory Job-Search-First Programs	SWIM	1985	31	x	Greenberg et al. (2010)
Welfare to Work San Diego Mandatory Mixed-Initial-Activity Programs	GAIN SD	1987	31	x	Freedman et al. (1996)
					Greenberg et al. (2010)
Welfare to Work San Diego Mandatory Work Experience Program	Work Exp. SD	1982	32	x	Brock et al. (1993)
					Greenberg et al. (2010)
Welfare to Work Tulane Mandatory Mixed-Initial-Activity Programs	GAIN Tul.	1988	31	x	Freedman et al. (1996)
					Greenberg et al. (2010)
Welfare to Work Vermont Earnings Supplement Programs	WRP Earn Supp.	1994	31	x	Greenberg et al. (2010)
					Scrivener et al. (2002)
Welfare to Work Vermont Time-Limit-Mix Programs	WRP Time lim.	1994	31	x	Greenberg et al. (2010)
					Scrivener et al. (2002)
Welfare to Work West Virginia Mandatory Work Experience Program	CWEP	1983	33	x	Brock et al. (1993)
					Greenberg et al. (2010)

Table II: MVPF, WTP and Cost Estimates with Confidence Intervals, All Programs

Program	MVPF	MVPF CI	WTP	WTP CI	Cost	Cost CI	Baseline
Child Education	∞	[12.14, ∞]	4.95	[3.06, 6.90]	-0.24	[-0.77, 0.26]	
Abecedarian	11.89	[-0.18, ∞]	2.62	[-0.24, 5.63]	0.22	[-0.76, 1.15]	x
CPC Extended	∞	[-∞, ∞]	4.15	[-21.80, 27.36]	-1.22	[-6.13, 4.36]	
CPC Preschool	∞	[∞, ∞]	2.23	[0.24, 4.21]	-0.35	[-0.62, -0.10]	
CPC School	1.32	[-∞, ∞]	-0.18	[-1.23, 0.83]	-0.14	[-1.27, 0.98]	
Head Start Geo	0.71	[0.00, ∞]	0.72	[-11.66, 10.22]	1.01	[-0.02, 1.66]	
Head Start RCT	2.41	[1.90, 3.15]	1.29	[1.10, 1.50]	0.54	[0.49, 0.59]	
K12 Spend	∞	[∞, ∞]	8.78	[4.58, 13.03]	-1.03	[-2.02, -0.06]	x
K12 Spend Mich.	0.65	[0.10, 2.48]	0.62	[0.01, 1.59]	0.95	[0.76, 1.08]	
Perry Preschool	43.61	[2.08, ∞]	3.45	[1.34, 5.75]	0.08	[-0.53, 0.64]	x
College Adult	-5.43	[-∞, ∞]	-2.68	[-124.88, 60.25]	0.49	[-8.18, 16.54]	
AOTC (IS)	6.75	[-1.51, ∞]*	2.45	[-6.98, 13.17]*	0.36	[-7.36, 6.86]*	x
AOTC (JE)	-1.77	[-29.91, ∞]*	-7.63	[-70.24, 45.78]*	4.31	[-12.57, 24.48]*	x
AOTC (JS)	∞	[-5.95, ∞]*	9.96	[-49.88, 74.47]*	-1.37	[-17.34, 11.92]*	x
AOTC (SI)	10.05	[-12.60, ∞]	5.36	[-84.56, 106.72]	0.53	[-13.07, 12.50]	x
AOTC (SE)	-0.02	[-2.32, ∞]*	-0.02	[-7.49, 6.32]*	1.12	[-6.35, 8.41]*	x
AOTC (SS)	∞	[-8.87, ∞]*	23.39	[-126.08, 187.02]*	-1.61	[-25.82, 18.34]*	x
HOPE Cred.	12.58	[-14.00, ∞]	5.27	[-758.34, 535.11]	0.42	[-63.63, 116.58]	x
HTC (IS)	18.86	[-2.83, ∞]*	5.91	[-26.04, 42.35]*	0.31	[-12.00, 11.91]*	x
HTC (JE)	2.37	[-2.18, ∞]*	8.21	[-39.98, 60.19]*	3.47	[-22.77, 28.43]*	x
HTC (JS)	∞	[-3.66, ∞]*	18.41	[-98.02, 144.89]*	-3.15	[-66.38, 57.71]*	x
HTC (SE)	11.83	[-4.69, ∞]*	4.59	[-19.34, 30.76]*	0.39	[-3.99, 3.89]*	x
HTC (SS)	-1.64	[-28.32, ∞]*	-1.91	[-23.38, 16.26]*	1.16	[-3.87, 6.69]*	x
HOPE/LLC	-8.81	[-∞, ∞]	-42.82	[-270.63, 9.61]	4.86	[-4.03, 34.72]	x
Adult Pell	2.18	[0.58, 7.55]	3.42	[1.00, 6.94]	1.57	[0.99, 2.24]	x
Tuition Deduc (JE)	0.77	[-1.80, 93.83]	1.00	[-4.52, 6.68]	1.29	[0.12, 2.46]	x
Tuition Deduc (JS)	-0.02	[-2.40, 4.61]	-0.03	[-5.40, 4.32]	1.38	[0.57, 2.46]	x
Tuition Deduc (SE)	∞	[-∞, ∞]	1.00	[-7.19, 8.79]	-1.13	[-1.99, -0.29]	x
Tuition Deduc (SS)	∞	[-∞, ∞]	5.38	[-1.70, 14.61]	-4.98	[-6.25, -3.18]	x
College Child	∞	[4.50, ∞]	8.79	[3.04, 15.49]	-0.38	[-1.77, 0.72]	
Cal Grant GPA	∞	[10.72, ∞]	9.41	[3.43, 16.44]	-0.57	[-1.63, 0.32]	x
Cal Grant Inc	-0.69	[-2.36, 7.41]	-1.04	[-5.37, 4.29]	1.51	[0.63, 2.21]	x
CUNY Pell	1.39	[-2.77, 12.14]	1.42	[-3.19, 7.25]	1.02	[0.54, 1.63]	x
CC Mich	29.46	[-2.70, ∞]	7.80	[-10.11, 29.91]	0.26	[-3.42, 2.84]	x
CC Texas	349.51	[2.16, ∞]	10.69	[2.35, 22.34]	0.03	[-2.20, 1.72]	x
DC Grant	22.98		7.62		0.33		x
FIU GPA	∞	[∞, ∞]	13.73	[1.40, 62.13]	-2.97	[-15.62, -0.02]	x
Florida Grant	7.42	[1.31, ∞]	7.40	[1.32, 16.06]	1.00	[-0.40, 2.27]	x
Free FAFSA (Dep)	4.03	[0.92, 10.89]	33.67	[3.24, 97.80]	8.35	[1.98, 20.10]	
Free FAFSA (Indep)	2.12	[-0.28, 12.66]	5.32	[-1.29, 16.40]	2.51	[0.72, 5.27]	
Georgia HOPE	4.00	[0.73, 22.55]	3.60	[1.11, 6.57]	0.90	[0.28, 1.48]	x
HAIL Aid	1.30	[0.29, 3.75]	0.97	[0.20, 1.80]	0.75	[0.55, 0.93]	
Kalamazoo	1.93	[1.00, 5.50]	1.93	[0.93, 3.79]	1.00	[0.79, 1.18]	x
MA Scholarship	0.72	[-1.02, 2.85]	1.21	[-2.09, 3.98]	1.68	[1.26, 2.22]	x
Ohio Pell	2.49	[1.00, 5.52]	2.88	[1.52, 4.47]	1.16	[0.80, 1.48]	x
TN Pell	0.77	[-1.26, 3.45]	0.78	[-1.48, 2.96]	1.02	[0.77, 1.28]	x
Texas Pell	∞	[∞, ∞]	85.74	[0.77, 173.60]	-17.38	[-33.15, -1.92]	x
Soc Sec College	7.26	[1.20, ∞]	5.03	[0.82, 12.38]	0.69	[-0.40, 1.26]	x
College Spend	4.00	[1.51, 19.65]	3.17	[1.44, 5.63]	0.79	[0.33, 1.13]	x
TN Hope	1.94	[1.04, 5.42]	1.94	[0.95, 3.58]	1.00	[0.77, 1.23]	x
College Tuition	1.02	[-1.02, 5.60]	1.02	[-1.35, 3.57]	1.00	[0.68, 1.29]	x
WI Scholarship	1.43	[1.06, 2.76]	1.46	[1.05, 2.19]	1.02	[0.92, 1.11]	x
Job Training	0.44	[-36.19, 0.90]	0.36	[-0.82, 1.51]	0.83	[-0.09, 1.75]	
Job Corps	0.15	[-0.23, 0.58]	0.15	[-0.23, 0.55]	0.98	[0.93, 1.03]	x
JTPA Adult	1.38	[-0.38, 2.09]*	1.17	[-0.17, 2.64]*	0.85	[0.08, 1.65]*	x
JTPA Youth	-0.23	[-3.72, 1.24]*	-0.21	[-1.70, 1.29]*	0.91	[0.15, 1.66]*	x
JobStart	0.20	[0.04, 0.42]	0.20	[0.06, 0.34]	1.02	[0.80, 1.24]	x
NSW Women	1.48	[-∞, ∞]	0.57	[-0.50, 1.64]	0.39	[-0.09, 0.86]	x
NSW Ex-Addict	0.44		0.35		0.79		x
NSW Ex-Offender	0.64		0.53		0.82		x
NSW Youth	0.60	[-∞, ∞]	0.47	[-4.23, 5.15]	0.78	[-3.32, 4.87]	x
Work Advance	0.78	[0.26, 1.34]*	0.64	[0.21, 1.11]*	0.83	[0.83, 0.83]*	x
Year Up	0.43	[0.37, 0.48]	0.41	[0.36, 0.45]	0.96	[0.95, 0.97]	x
Disability Ins.	0.85	[0.82, 0.88]	1.00	[1.00, 1.00]	1.18	[1.14, 1.22]	
DI Generosity	0.96	[0.95, 0.97]	1.00		1.04	[1.03, 1.05]	x

DI Judge	0.78	[0.72, 0.85]	1.00		1.28	[1.18, 1.39]	x
DI Examiner	0.74	[0.71, 0.78]	1.00		1.34	[1.28, 1.41]	x
DI Veterans	0.95	[0.92, 0.98]	1.00		1.05	[1.02, 1.08]	x
Health Adult	0.89	[0.56, 1.57]	1.49	[1.00, 1.99]	1.67	[1.01, 2.39]	
Mass HI (150%FPL)	0.80		1.00		1.25		x
Mass HI (200%FPL)	0.85		1.00		1.18		x
Mass HI (250%FPL)	1.09		1.00		0.92		x
Medicare Intro	1.63	[0.52, 3.83]	2.00	[0.58, 3.44]	1.23	[0.48, 1.78]	x
Oregon Health	1.16	[1.08, 1.25]	1.46	[1.19, 1.83]	1.26	[1.04, 1.57]	x
Medigap Tax	0.40	[0.22, 1.54]	1.00		2.53	[0.64, 4.44]	x
Health Child	∞	[25.05, ∞]	6.17	[3.09, 13.39]	-0.78	[-2.52, 0.17]	
MC Child 83+	∞	[0.26, ∞]	0.86	[0.66, 1.44]	-0.20	[-0.47, 1.82]	x
MC Pregnant & Infants	∞	[∞, ∞]	13.92	[6.03, 41.64]	-2.02	[-7.85, -0.27]	x
MC Child (State Exp)	∞	[-0.36, ∞]	8.13	[-0.24, 13.99]	-1.08	[-2.26, 0.57]	x
MC Intro	10.24	[0.93, ∞]	1.78		0.17	[-1.60, 1.93]	x
Supp. Sec. Inc.	0.75	[0.64, 0.85]	1.00	[1.00, 1.00]	1.33	[1.17, 1.56]	
SSI Review	0.76	[0.56, 1.00]	1.00		1.32	[1.00, 1.78]	x
SSI Judge	0.74	[0.72, 0.77]	1.00		1.34	[1.30, 1.39]	x
Unemp. Ins.	0.61	[0.53, 0.74]	1.20	[1.15, 1.24]	1.95	[1.63, 2.26]	
UI Ben (State Max)	0.68	[0.48, 1.13]	1.17	[1.11, 1.22]	1.71	[0.99, 2.41]	x
UI Ben (DD)	0.43	[0.28, 0.78]	1.17	[1.11, 1.22]	2.74	[1.51, 4.17]	x
UI Ben (DD w UR)	0.48	[0.30, 1.69]	1.17	[1.11, 1.22]	2.43	[0.79, 3.90]	x
UI Ben (GA)	1.03	[0.97, 1.09]	1.17	[1.11, 1.22]	1.14	[1.09, 1.18]	x
UI Ben (MO Exp)	0.74	[0.67, 0.81]	1.17	[1.11, 1.22]	1.59	[1.46, 1.73]	x
UI Ben (MO Rec)	0.44	[0.39, 0.50]	1.17	[1.11, 1.22]	2.68	[2.36, 3.01]	x
UI Ben (NY)	0.89	[0.82, 0.97]	1.17	[1.11, 1.22]	1.31	[1.21, 1.41]	x
UI Ben (RK)	0.84	[0.76, 0.92]	1.17	[1.11, 1.22]	1.40	[1.28, 1.52]	x
UI Dur (DD)	0.45	[0.25, 2.12]	1.30	[1.24, 1.36]	2.89	[0.61, 5.19]	x
UI Dur (MO)	0.83	[0.76, 0.90]	1.30	[1.24, 1.36]	1.57	[1.46, 1.69]	x
Housing Vouchers	0.70	[0.67, 0.74]	0.83	[0.82, 0.84]	1.18	[1.13, 1.23]	
HCV RCT to Welfare	0.76	[0.71, 0.81]	0.82	[0.80, 0.84]	1.08	[1.03, 1.13]	x
HCV Chicago Lottery	0.65	[0.61, 0.70]	0.83		1.27	[1.18, 1.37]	x
Jobs+	1.42	[0.45, 2.83]*	1.14	[0.41, 1.91]*	0.81	[0.67, 0.93]*	
MTO							
MTO	∞	[-2.80, ∞]	18.40	[-15.46, 51.85]	-2.44	[-11.35, 6.81]	x
Nutrition	0.61	[-1.82, ∞]	1.05	[-0.72, 2.78]	1.71	[-8.27, 7.47]	
WIC	1.38	[1.10, 1.66]	1.28	[1.08, 1.47]	0.93	[0.88, 0.98]	
SNAP Assist	0.92	[0.91, 0.96]*	0.92	[0.91, 0.96]*	1.00		x
SNAP Info	0.89	[0.89, 0.89]*	0.89	[0.89, 0.89]*	1.00		x
SNAP Intro	1.04	[-0.97, ∞]	1.09	[-2.45, 4.55]	1.05	[-0.38, 2.51]	x
SNAP Imm.	0.42	[-3.45, 1.00]	1.00		2.37	[-17.39, 13.41]	x
Cash Transfers	0.79	[0.47, 1.26]	0.88	[0.63, 1.20]	1.12	[0.94, 1.27]	
EITC 1986	1.20	[1.05, 1.38]	1.00		0.84	[0.73, 0.95]	x
EITC 1993	1.12	[0.27, 1.20]	1.00		0.89	[0.67, 1.18]	x
AFDC Generosity	0.87	[0.85, 0.90]	1.00		1.15	[1.11, 1.18]	x
AFDC Term Limits	0.81	[0.73, 0.90]	1.00		1.23	[1.11, 1.38]	x
Alaska UBI	0.92	[0.89, 0.96]	1.00		1.09	[1.05, 1.12]	x
Paycheck+	1.00	[0.87, 1.19]	1.00		1.00	[0.85, 1.15]	x
Neg Inc Tax	0.11	[-0.66, 3.81]	0.18	[-1.56, 2.39]	1.63	[0.49, 2.52]	x
Top Taxes	3.03	[1.35, ∞]	1.00	[1.00, 1.00]	0.33	[-0.09, 0.74]	
Top Tax 2013	1.16	[0.87, 1.92]	1.00		0.86	[0.54, 1.16]	x
Top Tax 1993	1.85	[1.21, 4.03]	1.00		0.54	[0.25, 0.82]	x
Top Tax 1986	44.27	[2.37, ∞]	1.00		0.02	[-0.37, 0.42]	x
Top Tax 2001	1.37	[0.92, 2.86]	1.00		0.73	[0.36, 1.09]	x
Top Tax 1981	∞	[0.94, ∞]	1.00		-0.51	[-2.13, 1.06]	x

Notes: This table presents our baseline estimates for each program in our extended sample, along the category averages reported in the bold header rows in each category. We exclude the welfare to work policies discussed in Section VI.C. For each policy, we report its MVPF, cost per dollar of programmatic spending, and willingness to pay per dollar of programmatic spending. We also report bootstrapped 95% confidence with adjustments discussed in Appendix H. The final column indicates whether the program is included in the baseline estimates (and thus included in the category averages). Confidence intervals are marked with a "*" in cases where we infer p-values using reported interval ranges. Programs in which the confidence interval is either inferred from p-values or missing are excluded from category averages.

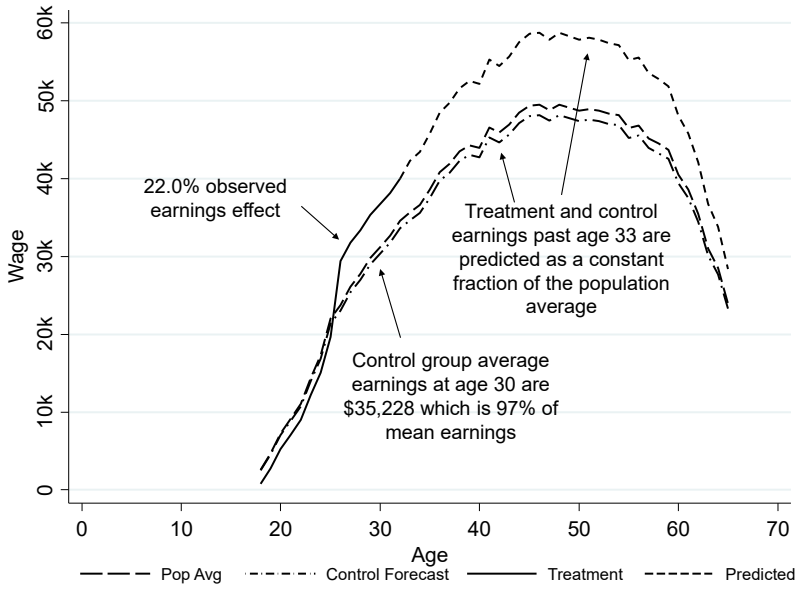
Table III: Publication Bias Estimation

	Baseline Sample						Restricted Sample					
	Children Estimates			**Adult Estimates**			**Children Estimates**			**Adult Estimates**		
Z-Score	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Z > 1.64	3.69	-		0.25	-		3.72	-		2.36	-	
	(1.49)			(0.30)			(1.58)			(1.38)		
Z < -1.64	1.14	-		0.22	-		1.07	-		8.44	-	
	(0.36)			(0.31)			(0.36)			(5.89)		
Z [1.64,1.96]			3.48			0.87			3.50			1.35
			(1.47)			(0.69)			(2.17)			(1.38)
Z [-1.96, -1.64]			1.01			1.13			0.99			4.26
			(0.45)			(0.54)			(0.43)			(2.57)
Z > 1.96	-	3.24	3.90	-	0.14	0.14	-	3.25	3.94	-	3.01	3.29
		(1.26)	(1.05)		(0.19)	(0.19)		(2.67)	(2.57)		(1.54)	(3.10)
Z < -1.96	-	1.21	1.23	-	0.10	0.11	-	1.12	1.13	-	11.23	12.48
		(0.49)	(0.48)		(0.16)	(0.17)		(0.57)	(0.46)		(3.93)	(3.19)
N	237	237	237	150	150	150	233	233	233	146	146	146

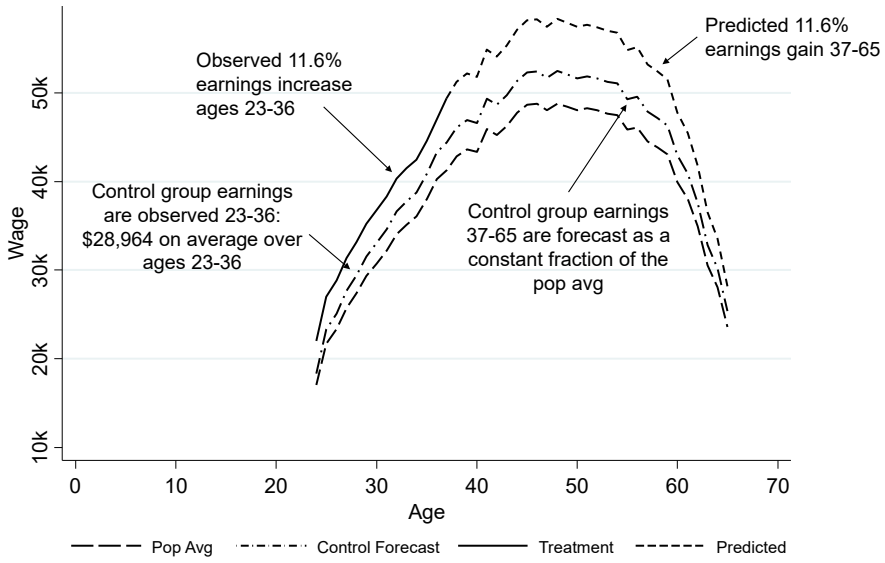
Notes: The numbers shown are the estimated probability of publication relative to an insignificant result. Standard errors in

APPENDIX FIGURE I: Income Projections Using the ACS

A. Predicted Earnings Impacts by Age for Florida GPA (Zimmerman, 2014)

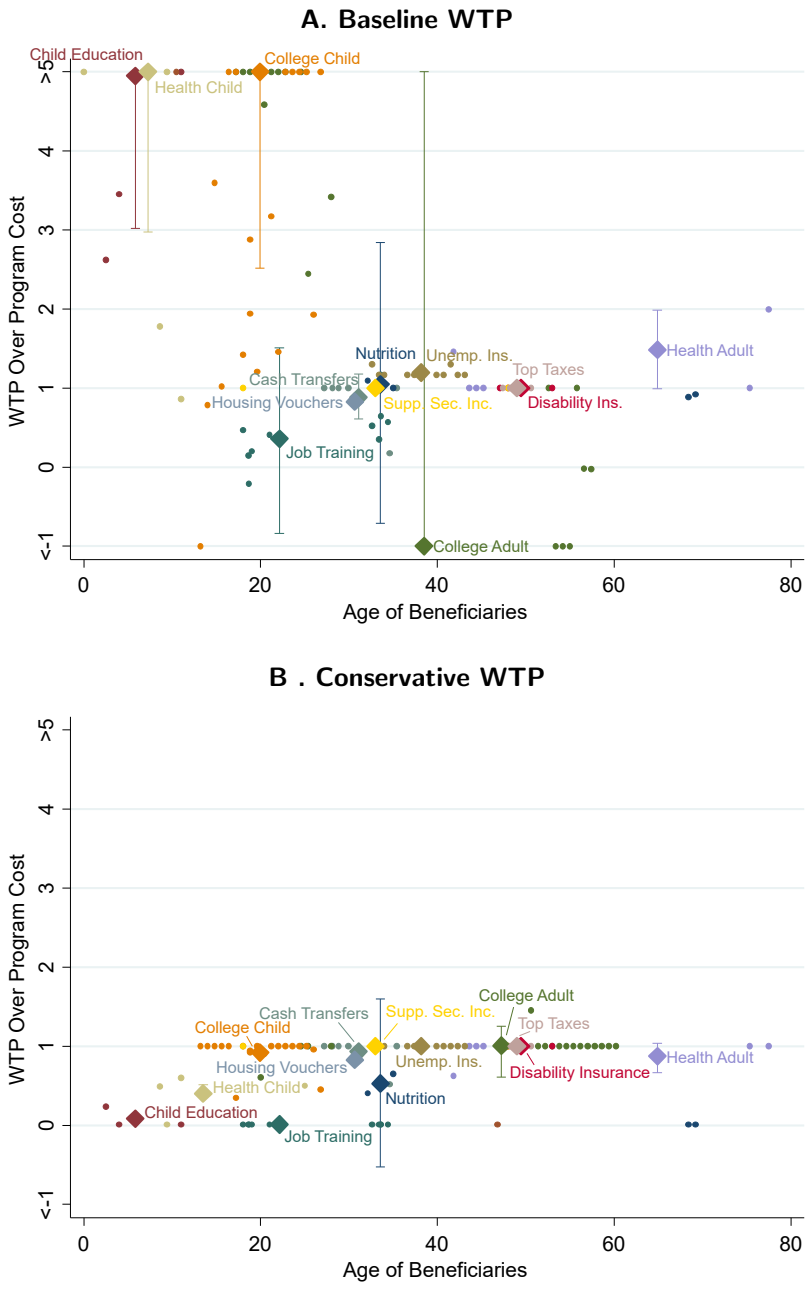


B. Predicted Earnings Impacts by Age for Medicaid Pregnant Women & Infants (Miller and Wherry, 2018)



Notes: Panels A and B present a decomposition of the elements which make up our income projection process for the examples in Section III.A. The "Pop Avg" series is constructed in each case from the 2015 ACS and using a 0.5% wage growth assumption. At each age "Pop Avg" gives the mean wage level that would prevail in the population for individuals of that age, when individuals in the treatment group for the relevant policy were that age. This number is constructed by assuming that the mean wage level at each age will rise (and has previously risen) by 0.5% in each year. The "Control Forecast" series is constructed by taking an estimate of earnings for a relevant control group at a particular age or range of ages, then calculating the implied proportion of the "Pop Avg" series at those ages, then projecting the series forwards (and backwards) as this constant fraction of "Pop Avg". The "Treatment" series is constructed by summing the observed treatment effects in dollar terms and the "Control Forecast" series. To construct the "Predicted" series we take the final value of the "Treatment" series, then calculate the ratio of this value to the value of the "Pop Avg" series at that same age, before applying this ratio to the "Pop Avg" series up to age 65. See Appendix I for further details of this methodology.

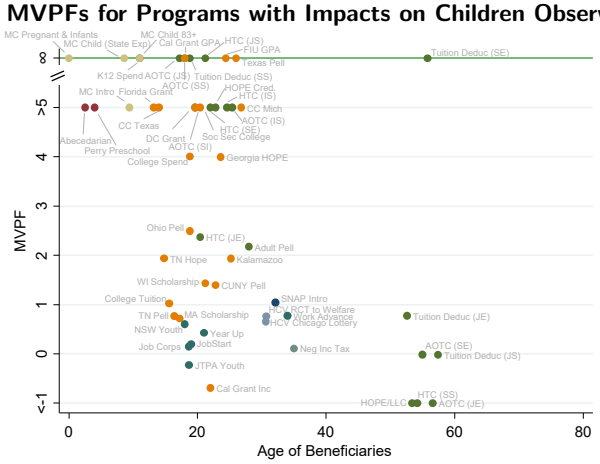
APPENDIX FIGURE II: Willingness to Pay per Dollar of Programmatic Spending



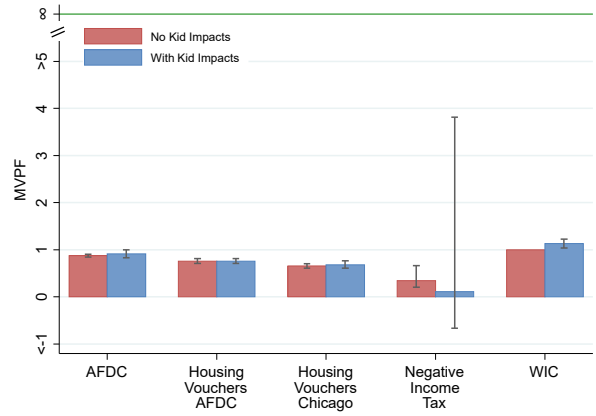
Notes: This figure presents estimates of WTP normalized by initial programmatic spending for each category-average group of policies in our baseline sample. Panel A presents estimates for our baseline WTP estimates. Panel B presents estimates from our conservative WTP method. We plot these estimates as a function of the average age of each policy's beneficiaries within category. Bootstrapped 95% confidence intervals with adjustments discussed in Appendix H are shown for the category averages. The normalized willingness to pay of individual policies are shown in smaller dots, color-coded to align with their respective categories.

APPENDIX FIGURE III: Robustness to Child Effects

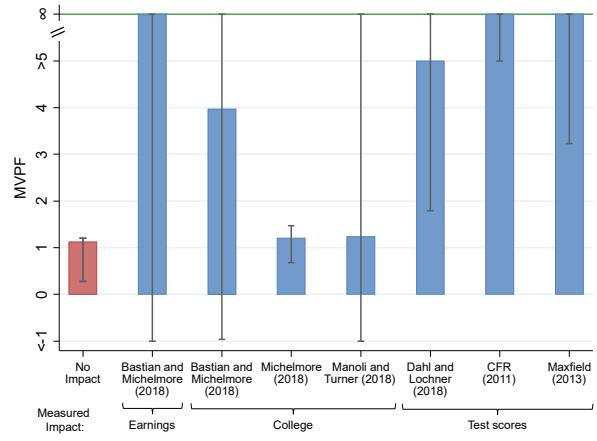
A. MVPFs for Programs with Impacts on Children Observed



B. Effects of Incorporating Impacts on Children

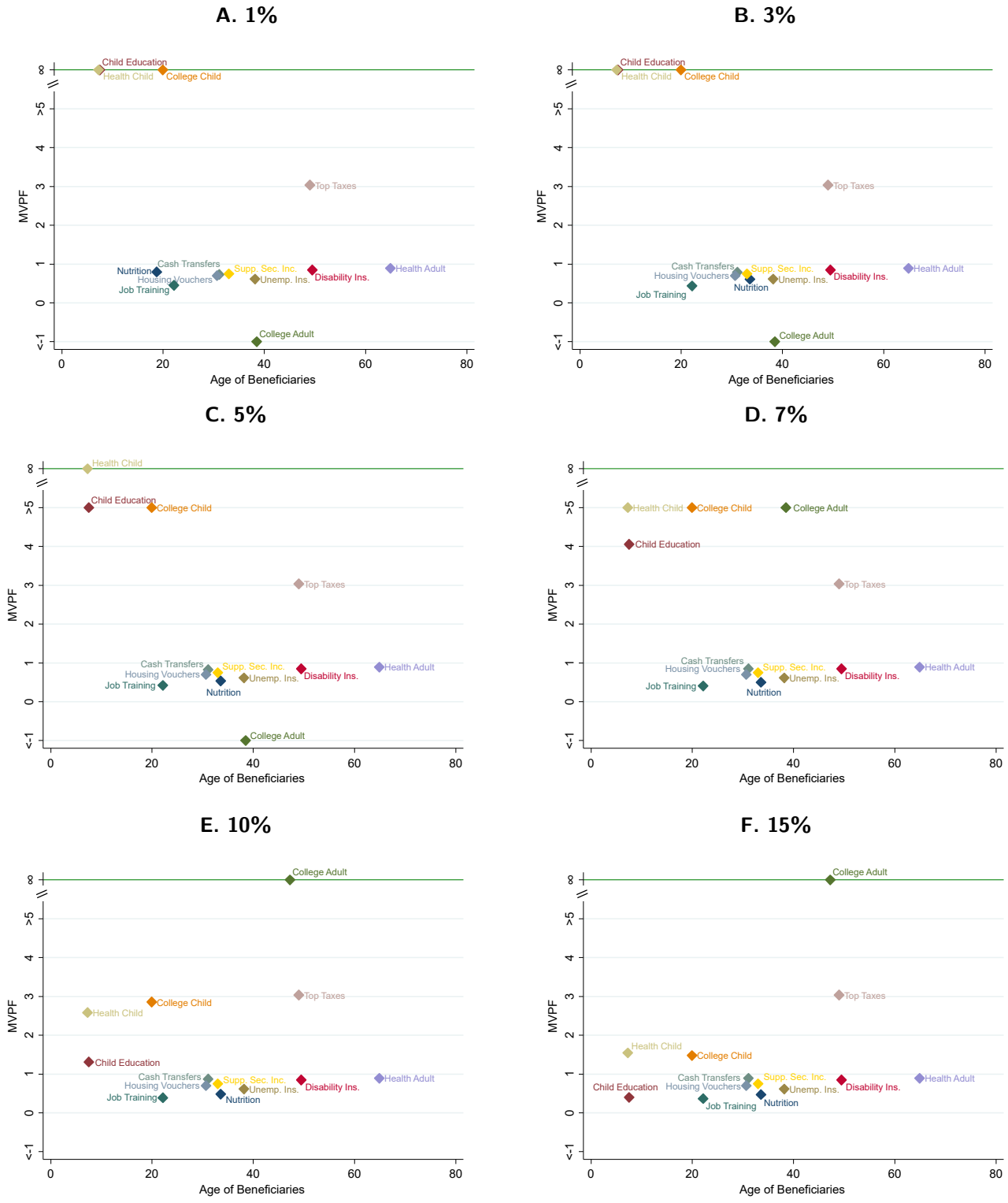


C. Inferred Estimates of the Impact of the 1993 EITC Reform on Children



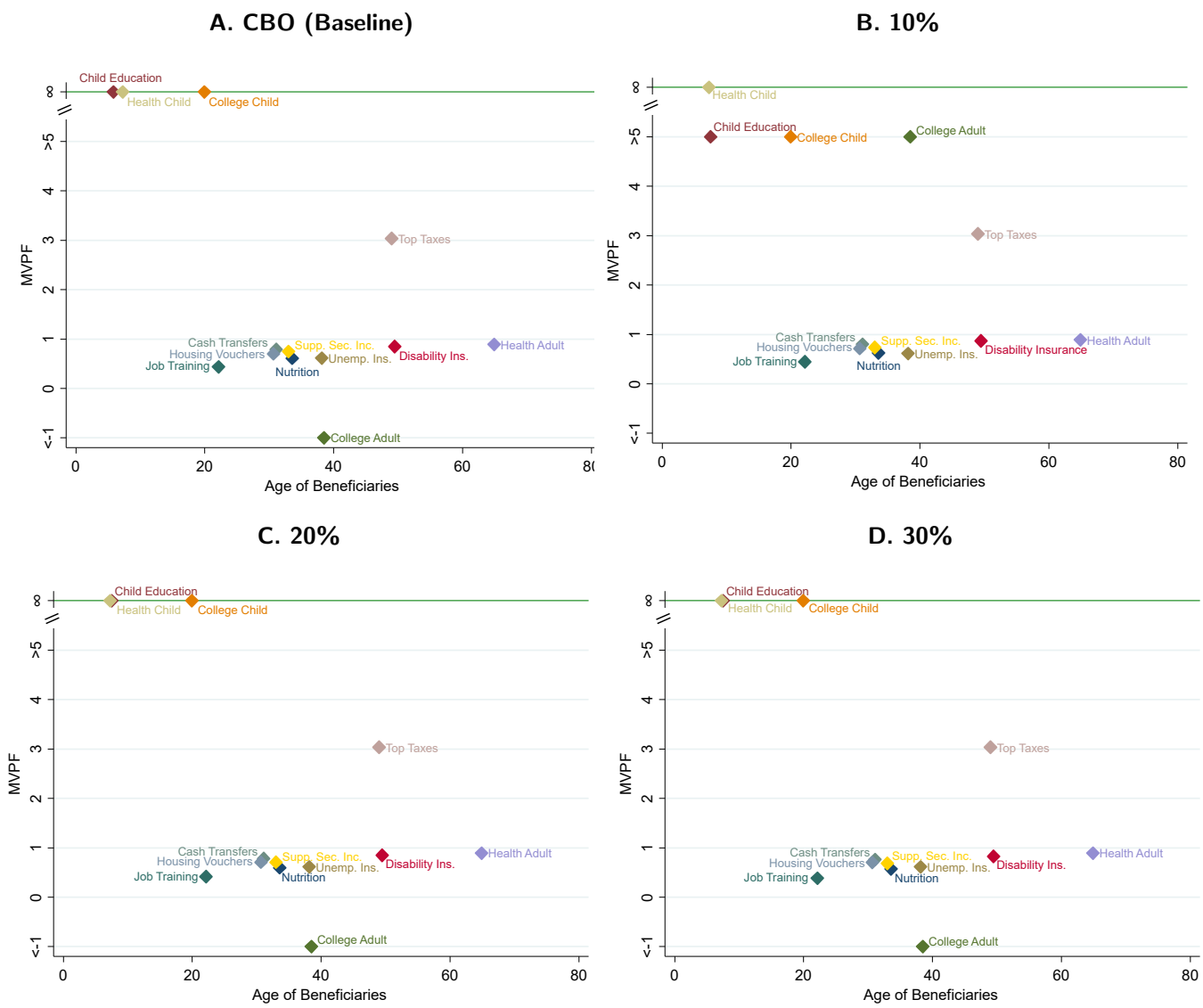
Notes: This figure assesses the impact of observing child impacts on our estimates as a function of the average age of the economic beneficiaries of the policy. Panel A restricts our sample to the subset of policies for which we observe estimates of the impact of the policy on children. In addition, panel B shows projected MVPFs for additional policies that do not observe earnings impacts but do observe another intermediate outcome such as birthweight (AFDC), college attendance (Housing Vouchers to AFDC Recipients), test scores (Housing vouchers in Chicago), and income (the introduction of food stamps, SNAP). Panel C reports the MVPF for the EITC under alternative methods of incorporating indirect impacts on children through test scores, college attendance, and income of EITC more broadly.

APPENDIX FIGURE IV: Robustness to Alternative Interest Rates



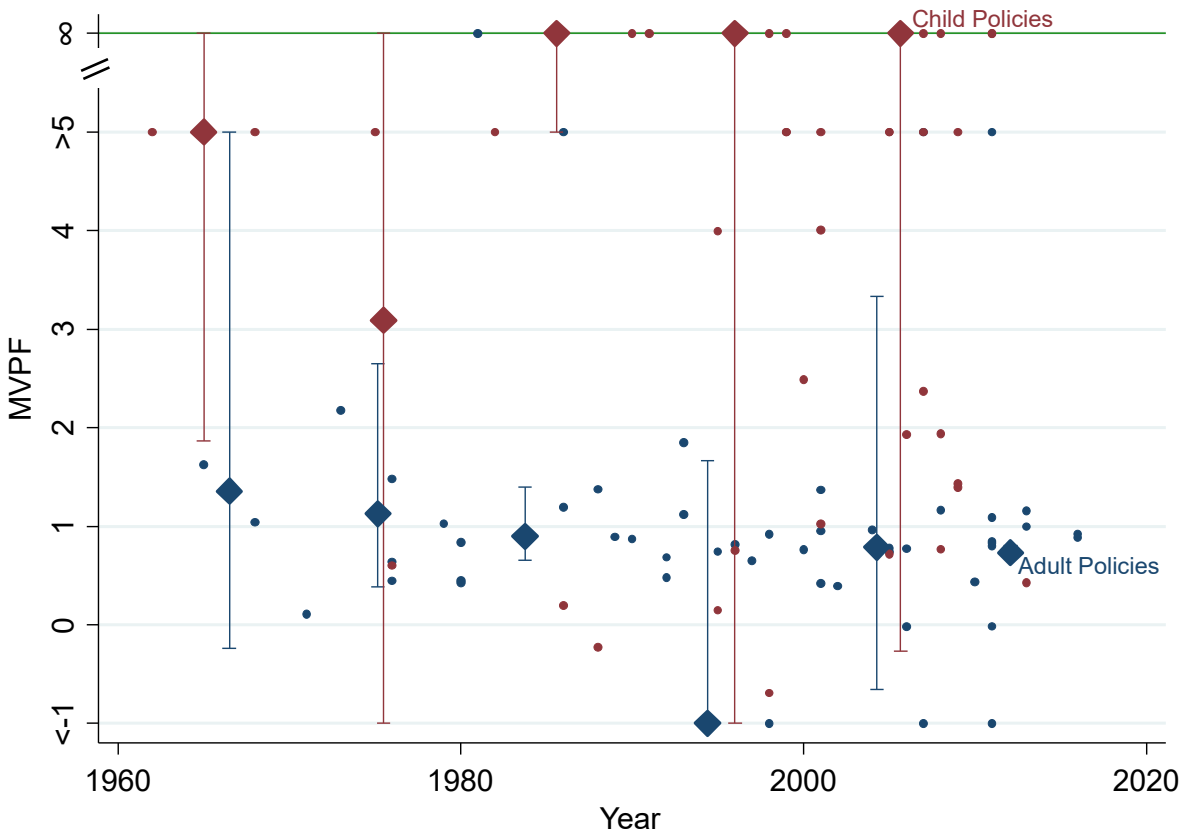
Notes: This figure presents our MVPF estimates as in Figure III and the category averages as in Figure IVB under alternative real interest rate assumptions, as opposed to our baseline specification of 3%. Panel B differs slightly from our baseline specification because we restrict to the subset of policies for which we are able to vary the discount rate (e.g. we exclude papers where we directly import an MVPF that relied on a particular discount rate). We omit confidence intervals for ease of viewing, but caution the reader that the estimate for the College Adult category has a CI that includes 0 and infinity.

APPENDIX FIGURE V: Robustness to Alternative Tax Rates



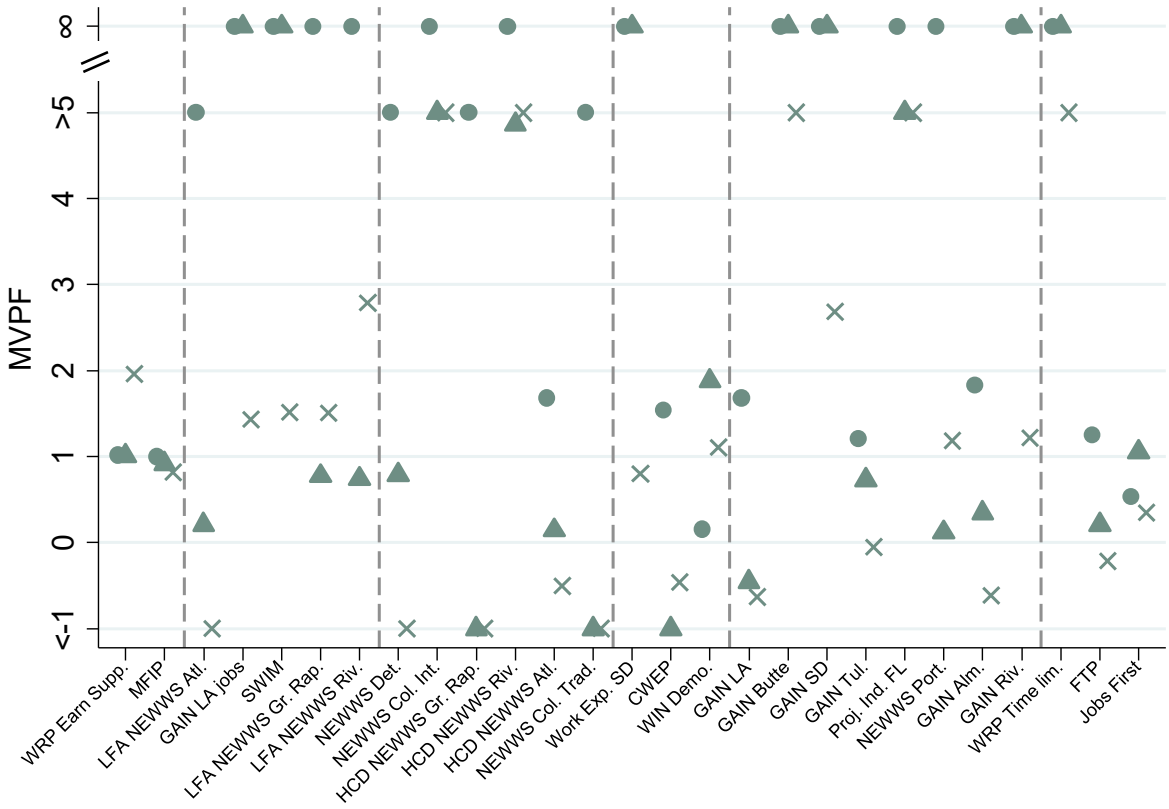
Notes: This figure presents our MVPF estimates as in Figure III and the category averages as in Figure IVB under alternative tax rate assumptions. Panel A replicates our baseline specification using the CBO estimates of the tax rates. Panels B-D adjust the tax rate to 10%, 20%, and 30%.

APPENDIX FIGURE VI: MVPFs by Decade



Notes: This figure presents MVPF for all policies evaluated in the paper based on the year in which the policy was implemented. Policies are divided into categories based on their decade of implementation and the average age of their economic beneficiaries. For policies implemented in each decade there are two categories – policies with beneficiaries over age 23 and policies with beneficiaries aged 23 or younger. Within each decade by age category we construct the MVPF for a hypothetical policy that allocates $1 of programmatic spending equally amongst all the policies in the category. This is the same approach used to create MVPF estimates for policy domains in previous figures. The capped lines show the 95% bootstrapped confidence intervals with adjustments discussed in Appendix H.

APPENDIX FIGURE VII: Welfare Reform



Notes: This figure presents estimates of the MVPF of 26 welfare reform policies discussed in Section VI.B. We report MVPF estimates using three potential measures of WTP: (1) Cost, the mechanical cost of the program incurred by the government, excluding any fiscal externalities from behavior change. Estimates from this specification are denoted by circles. (2) Change in transfer payments (welfare, food stamps and Medicaid). Estimates from this specification are denoted by Xs. (3) Change in post-tax income, which includes the change in participants' incomes due to changes in employment, and the change in their transfer payments. Estimates from this specification are denoted by triangles.