The Impact of Government Spending Shocks: Evidence on the Multiplier from State Pension Plan Returns

By Daniel Shoag*

shoag@fas.harvard.edu

JOB MARKET PAPER

Abstract

The effect of government spending on income and employment is a central unresolved question in macroeconomics. This paper employs a novel identification strategy to isolate exogenous and unexpected variation in state government spending. State governments manage large defined-benefit pension plans for which they bear the investment risk. Using a newly-collected dataset on the returns and portfolios of these plans, I show that the idiosyncratic component of their returns is a strong predictor of subsequent state government spending. Instrumenting with this 'windfall' component of returns, I find that state government spending has a large positive effect on income and employment. Baseline estimates indicate that each dollar of spending raises in-state income by $2.12, and that $35,000 of spending generates one additional job. These effects are not due to in-state investment bias, are concentrated in the non-traded sector, and are larger during times of labor force 'slack.' Finally, I consider how these results compare with the predictions of a standard macroeconomic model and outline which features in the model are consistent with the empirical findings.

*I gratefully acknowledge the patient guidance of my adviser Emmanuel Farhi, and my committee members Raj Chetty, Matthew Weinzierl, David Laibson and Robert Barro. I have benefited tremendously from their advice. I also want to thank Erik Hurst, Gita Gopinath, Alp Simsek, Benjamin Friedman, Greg Mankiw, Marios Angeletos, Alberto Alesina, John Campbell, Lauren Cohen, Chris Malloy, Daniel Bergstresser, Julio Rotemberg, and David Mericle for their suggestions. Yossi Ciment provided excellent research assistance. This research was supported by a NSF IGERT grant, the Tobin Project, the Taubman Center and the Center for American Political Studies.

1 Introduction

Does government spending increase income and employment? This question is central to macroeconomics and important to policymakers. Persuasive empirical evidence on this issue has been limited, however, by the difficulty of identifying shocks to government spending that are exogenous, unexpected, and large in magnitude.[1]

This paper introduces an instrumental variables approach that addresses these problems. U.S. states administer large, funded, defined-benefit pension plans and invest these funds differently. The defined-benefit nature of these plans makes the state government, not the pensioners, the residual claimant on the investment returns. Using newly assembled data, I isolate an exogenous 'windfall' component of these returns. I demonstrate that the idiosyncratic windfall returns earned by these plans are powerful predictors of future state-level spending. These windfalls can be used to identify random variation in spending at distinguishable dates.

This paper's research design can be clarified by the following thought experiment. Suppose New Hampshire's retirement system outperforms Tennessee's system after controlling for asset allocation. This performance is not due to the relative economic conditions in New Hampshire and Tennessee. Moreover, this excess return is presumably not expected ahead of time, or Tennessee would have altered its investment strategy. In the years after this outperformance, however, New Hampshire will have a smaller actuarially-required contribution to its pension plan and more resources available for government purchases. This paper estimates the effect of this randomly assigned spending on income and employment.

To calculate these windfalls, I constructed a new dataset on state pension plan asset holdings and investment returns. These plans are of first-order economic importance in their own right, with 19 million members and more than \$2.3 trillion in assets. There is tremendous public concern about their funding status and their impact on state budgets (Rauh and Novy-Marx, 2010). Despite this, there is little data available on their investment strategies and performance. Using open records requests and audited financial reports from the Library of Congress and various state libraries, I collected data for 83 systems over 21 years. The data reveal substantial variation in returns across plans, with an average in-year standard deviation of 300 basis points. These data are the most comprehensive available on state-administered plans and allow me to build a large panel of windfall shocks.

[1] For a detailed discussion of how omitting expectations can bias empirical work on government spending shocks, see Ramey (2009).

Pension windfalls are an attractive instrument for a number of reasons.[2] In addition to being exogenous and unexpected, pension return shocks vary in both the cross-section and the time-series. The identification procedure, therefore, does not fundamentally rely on strong assumptions about the comparability of different states and years. Additionally, the enormous resources managed by state-administered plans and the substantial volatility of returns generate large idiosyncratic shocks. The size of these shocks makes identification possible despite the natural variability in the outcome variables.

Using idiosyncratic pension returns as an instrument, my baseline estimates indicate that each dollar of government spending generates $2.12 of personal income. Baseline estimates for employment indicate that $35,000 of additional spending generates another job in state. These effects are stronger when employment and labor force participation are low, are present across state borders, and are concentrated in non-tradable industries. Both the first- and second-stage results are shown to be highly robust across specifications, controls, and constructions of the instrument.

These findings are puzzling in the context of the standard frictionless macroeconomic model. I explore this issue in a simple open-economy framework and demonstrate that this model is unable to generate an increase in employment and a multiplier greater than one. I then incorporate wedges into the original model in the form of nominal frictions. I discuss the role these frictions play in producing large multipliers and consider the implications these types of frictions would have for multipliers in other models.

My work differs from the traditional multipliers estimated in the literature in that I explore the effects of windfall-driven state-level spending. The macroeconomic literature on government spending multipliers, however, typically uses national or international data and VAR techniques. This literature (key papers include Blanchard and Perotti 2002 and Mountford and Uhlig 2005) identifies government spending shocks by imposing time or sign-restrictions on the impulse responses. Unfortunately, the estimates in this literature appear highly sensitive to the identification assumptions used and the choice of countries, spending definitions, time-frame, variables, and lag lengths included in the VAR (Perotti 2004, Pappa 2010). Additionally, this method is not able to control for the expectations of spending, nor is it easy to map the recovered innovations to narrative accounts of fiscal policy.

A second strand of the literature (recent papers include Ramey 2010 and Barro and Redlick 2010) estimates multipliers using military dates identified from narrative accounts. This approach has the advantage of isolating clear dates at which the expectations of future government spending change. There is a question, though, of the extent to which these shocks isolate variation in spending from simultaneous policy

[2]This identification procedure is similar to the one used by Rauh (2006) to identify financial resource shocks for corporations.

changes and economic conditions. This procedure is also limited to a small number of events and an atypical composition of spending.

There is a nascent IV literature that instruments for government spending at the state level.[3] Clemens and Miran (2010) use variation across states in balanced budget amendments to estimate the income effects associated with spending rescissions. Chodorow-Reich, Feiveson, Liscow and Woolston (2010) use the variation across states in 2007 Medicaid expenditure to instrument for ARRA grants. Fishback and Kachanovskaya (2010) instrument for New Deal spending at the state level using the political competitiveness of different states. Though these papers use very different sources of identification, their baseline results are similar to the ones found in this paper.[4]

An important caveat is that this paper does not estimate the effect of a national increase in spending funded by an increase in taxes or the issuing of debt. The exogenous movements in spending identified in this paper are funded, in a sense, by the 'windfall' return. To the extent that there are large tax or interest rate effects, an unfunded multiplier may be substantially different than the results presented here. Additionally, economic spillovers and the migration of resources are likely to be larger across states than across countries. In light of these differences, the estimates in this paper are informative about the effectiveness of national fiscal stimulus only indirectly.

This caveat does not imply that the income and employment effects recovered here are irrelevant for policy. The windfalls in this paper are relative windfalls, meaning that they provide 'free' resources to the state but do not represent an increase in the aggregate resources produced. As such, these windfalls closely resemble cross-state transfers and are directly informative about the ability of these transfers to stimulate a region. State-level risk within the United States is substantial,[5] and policymakers might be interested in the feasibility of smoothing this risk using transfers to state governments. Thus, while this paper is informative about national stimulus indirectly, it provides a direct estimate of the effectiveness of fiscal federalism for regional risk-sharing.

This paper proceeds as follows: Section two provides some background information on state-administered pension funds and their funding process and describes the construction of the data set. Section three explores threats to the validity of the identification strategy, focusing on the possibility of in-state investment bias. I

[3]Cohen, Coval and Malloy (2010) instrument for spending at the state-level, but do not explore the effects on income or employment. Kraay (2010) instruments for spending in developing countries. An earlier literature looked at the employment effects of spending using cross-state variation in military expenditure. Hooker and Knetter (1997) find that this type of spending has large employment effects.

[4]Clemens and Miran report a baseline multiplier of 1.7 and Chodorow-Reich et al. report that $28,500 in grants generates an additional job. Fishback and Kachanovskaya report a multiplier on public works and relief of 1.67.

[5]Athanasoulis and van Wincoop (1998) estimate that the 95% confidence interval of 5-year individual state GDP growth rates ranges from -7% to 25%, after controlling for the realized national growth rate.

discuss the sources of variation in pension fund returns and the construction of the instrument in section four. Section five demonstrates the robust first-stage relationship between state-administered pension windfalls and state government spending. It also demonstrates the lack of a relationship between these windfalls and state tax rates. This motivates using these windfall gains as an instrument for spending alone. Section six presents the baseline instrumental variable results for income and employment, as well as several robustness tests. Section seven explores the mechanisms by which increased government spending raises income and shows that the effect is stronger in non-tradable industries and when economic slack is high. Section eight compares the empirical findings with those generated in a conventional macroeconomic model and explores which features are needed to match the data. Finally, I conclude with a discussion of these findings in the context of the literature and their potential policy implications.

2 State-Administered Pension Plans: Background and Data

Roughly one in seven American employees works for a state or local government, and almost 80% of these employees are covered by a public pension plan.[6] During 2008, state-administered plans held $2.3 trillion in assets, had 19 million members, and owned 7.9% of total U.S. corporate equities.[7] These plans are overwhelmingly defined-benefit. In 2008, 97% of public pension plan assets were held in DB plans. Benefits accrued by state-administered pension plans are protected under most state constitutions, and public pension liabilities have been honored even in bankruptcy or extreme fiscal stress.[8] State plans are predominantly pre-funded, and state contributions to these plans in 2007 ranged from 0.3 to 5.7 percent of state general expenditure. Figure 1 displays the distribution of pension fund asset holdings relative to state government revenue.

From the above statistics it is clear that state-administered pension plans hold a large fraction of America's liquid financial assets and represent a major component of state government activity. Moreover, given the large sums involved (asset holdings on the order of $10,000 per person nationally), the investment outcomes of these plans have serious consequences for state budgets. Despite this, there is a surprising lack of data on the investment portfolios and outcomes for these pension plans.

[6]Munnell, Aubry and Muldoon (2008). In 2007, more than 90% of the membership of these plans and more than 80% of the assets belonged to state-run plans (Census Government Employee Retirement Systems Survey, 2007)

[7]Fed Funds Table Z1. For comparison, in 2008 the total financial assets of broker-dealers was $2.2 trillion. At the end of 2007 these figures were $3.2 trillion for state and local pension funds and $3.1 trillion for broker-dealers.

[8]Bovbjerg GAO report (2008). Neither New York City in the 1970s, nor Orange County when it declared bankruptcy in 1994, failed to honor DB pension commitments.

As part of this study, I have constructed a new data set containing fiscal year returns for more than 83 of the largest state administered retirement systems. Each data point was collected from an audited financial report or from a system's response to an Open Records Request. Financial reports were obtained from system or state websites, at the Library of Congress and at various state libraries. The data set includes at least one plan per state and runs from 1987-2008, with a 93% completeness rate.[9] This is the longest and most comprehensive data set on state-administered pension plan returns of which I am aware. A visual display of this data for plans with fiscal years ending in June is displayed in Figure 2.

The time series data reveal substantial variation in returns. The annual standard deviation among plans with fiscal years in June ranges from 1.6-3.9 percentage points. Over time, this variation compounds into major differences in investment performance. Using the plans with June fiscal years and complete records from 1987-2008, I find dramatic differences in the total cumulative returns earned by plans. The median cumulative return is 712% with a standard deviation of 94%. The bottom 5% of the sample earned below 530%, and 5% of the sample earned above 833%. Given the size of these plans, the difference between the top and bottom performers is on the order of three years of state expenditure.

My research design also requires data on the asset allocation and market value of these systems. The trade publication *Pensions and Investment* has conducted a yearly survey of the largest 200 systems, beginning in the early 1980s. Data for the majority of plans was generously provided by the publication or manually recorded from back issues of this biweekly. This data was augmented with information from the *Money Managers Directory of Pension Fund and their Managers* for a number of smaller plans. Once again, these data on the investment allocations of state-administered plans are the most comprehensive data of which I am aware.

Additional data on state-administered plans were also taken from the biennial PENDAT survey (Zorn 1991, 1993, 1995, 1997, 2001), the Center for Retirement Research at Boston College, the Wisconsin Legislative Council's Comparative Retirement Studies, NEA reports and the Public Funds Survey conducted by the National Association of State Retirement Administrators and the National Conference on Teacher Retirement (Brainard, 2007). A detailed description of these data sources, as well as the auxiliary sources for this project, can be found in the online appendix.

[9] In my empirical work, I drop Oregon and New Jersey. Oregon is dropped due to the very short time period for which data was available (2003-2008). New Jersey was dropped due to concerns about the validity of the Department of Pensions and Benefits financial records. The SEC recently sued New Jersey, alleging fraud in these reports. ("Pension Fraud by New Jersey is Cited by SEC", *NY Times* 8/18/2010). Including these states in the estimation does not change the central results. Including Oregon generates a baseline multiplier of 2.0 (relative to the current finding of 2.1). Including New Jersey generates a benchmark multiplier of 2.2, and including both generates a point estimate of 2.0.

Data on state and local government expenditure were taken from the Census Bureau's State Government Finances survey,[10] and data on state-level employment are taken from the BLS Current Employment Survey, Quarterly Census of Employment and Wages, Business Dynamics Survey and Local Area Unemployment Statistics. Data on state-level personal income are taken from the BEA Regional Income series.

The use of state personal income as the dependent variable, as opposed to gross state product, warrants some discussion. Unlike gross state product, data on personal income are available from the BEA at the quarterly frequency. This makes it possible to match income to the fiscal year periods for which state government spending is reported. Additionally, there is a break in 1997 in the method used by the BEA to calculate gross state product corresponding to the switch from SIC to NAICS industry classifications. The BEA counsels researchers to use the personal income series, rather than the GSP series, when using data pre- and post-1997. In accordance with this advice, this paper uses only personal income and not GSP. [11]

There are a number of differences in the measurement of personal income and GSP. The most important for the purposes of this paper is that personal income includes income from transfers. I explore this issue in a later section. Other differences include accounting for the consumption of fixed capital and cross-border payments. Interested readers are referred to the BEA website

2.1 Review of Funding Procedure and the Flypaper Effect

Before beginning, it is worth reviewing the standard funding mechanism used by state-run plans. The typical process begins with an actuarial valuation study, which analyzes investment outcomes, transfers in and out of pension funds, and any changes to the path of future liabilities. Based on these reviews, an actuarially required contribution (or ARC) is determined for the state government and/or other employers. The contribution is set so as to fund the ongoing incurred liability (the 'normal cost') and to amortize the system's unfunded liability. This ARC is a function of the system's investment returns in that these returns affect the funding status of the plan. Many systems smooth the impact of return shocks on required contributions. Typically the difference between the actual and assumed rate of return is realized over 0 to 5 years. Within this general framework there is a substantial amount of institutional heterogeneity. More details on the funding processes can be found in Appendix A.

[10] The definition of state government used by the Census differs somewhat from the definition used by the BEA in the construction of the national income accounts. Interested readers are referred to NIPA Table 3.19. Under the Census' definition, state government spending comprises 15-20% of U.S. GDP.

[11] In unreported regression, I do use the time series of GSP as the dependent variable of interest. The effects are similar to those reported for personal income, with somewhat higher multipliers (close to 3, depending on the specification). The estimates are less precise—presumably due to the measurement issues discussed above—and the statistical significance is inconsistent across specifications.

In theory, state governments are required to allot the full ARC, though this constraint binds with varying severity across states. Legally, the government in every state appears to have the ultimate authority over the contributions to the state-run plans.[12] Despite this, for every year in which data is available the majority of plans in my sample had a contribution equal to or above their ARC. Even among plans failing to make the ARC, the average percentage of the ARC contributed throughout the sample was 74%. The high contribution rate may stem from the political pressure of public sector unions, which often act as powerful watchdogs for these plans.[13] Figure 3 plots the average percentage of the ARC contributed over time for state-run plans.

Though state pension contributions are a function of their systems' investment returns, it is not immediate that the amount a government has to contribute ought to affect spending in other areas.[14] A rational benevolent government should choose the level of spending where the marginal benefit equals the marginal cost of raising funds. A rational government should not therefore spend a higher percentage of this windfall than it would spend of general income. Additionally one might expect state governments to smooth one-time windfalls over many years.

As I will demonstrate below, the data show that pension windfalls are associated with a substantial increase in spending in the short-run. The spending response from income earned by the state government is much larger than the average state tax rate and is concentrated in the first three years. This phenomenon, known as the 'flypaper effect' has been well studied in the literature (Gramlich 1977, Hines and Thaler 1995). This effect strengthens the relationship between windfall returns and spending, making it possible to identify large changes in spending.

3 Testing the Possible Endogeneity of Returns

Given the existence of the flypaper effect, the idiosyncratic returns earned by state pension funds will predict future spending. Before using these returns as an instrument for spending, though, it is important to establish that these returns are orthogonal to in-state economic shocks. One clear threat to this identification procedure is the possibility that state-administered plans might target their investments in state.

[12]There have been a number of lawsuits brought by public retirement systems against their administering governments for failing to make actuarially required contributions. The majority of these cases have been decided in favor of the government, ruling that while benefit payments were constitutionally protected, the timing of funding was not. (See Retired Public Employees Council of Washington v. Charles, 62 P.3d 470 (Wash. 2003) and People ex rel., Sklodowski v. State, 695 N.E.2d 374 (Ill. 1998)). One exception can be found in Firemen's Retirement System v. City of St. Louis, 2006 WL 2403955 (Mo. App. E.D. Aug 22, 2006).

[13]For example, in 2009 Ohio governor Ted Strickland submitted a proposal to reduce the state's pension contributions in the face of a state-wide budget crisis. Opposition from public sector unions killed the proposal.

[14]Note that contributions to the pension fund, which essentially move money from one account to another, are not considered spending

Overinvestment in state specific industries or companies would create a spurious correlation between a state government's expenditure and its subsequent economic performance. Establishing the degree of this bias is therefore a crucial first step to establishing the validity of my instrument.

A number of papers have looked at the issue of in-state bias or economically-targeted-investments (ETI's). A thorough survey by Munnell and Sundén (2000) concluded that, while heavily used by plans in the 1970s, ETI's now constitute a small and declining fraction of the assets held by state-administered plans. Munnell and Sundén use a 1995 GAO survey to place an upper bound on this type of activity of 2.4% of total public pension assets in that year. A more recent paper by Brown, Pollet and Weisbenner (2009) looks at the in-state-bias in the domestic equity investments of these plans from 1980-2008. They, too, find small biases in the range of 0.9% to 4.2% of the measured portfolio.

Like these earlier studies, I find that the in-state biases of these plans are too small to meaningfully drive the variation in my instrument. In this section, I present four types of evidence to that effect.

The simplest way of learning whether or not system administrators skew their investment in-state is to ask them. In Table 1, I present the results of a survey question administered as part of the biennial PENDAT survey conducted throughout the 1990s. The self reported degree of bias is extremely small across all of the years in the sample, with the vast majority of plans reporting no ETIs.

To verify the self-reported results, I collected data from the EDGAR database of SEC filings. 13F forms must be filed by institutional investment managers supervising more than \$100 million in 13(f) securities[15] typically domestic equities. The applicability of this law to public plans is unclear, yet a number of plans do file quarterly with the SEC. These filings contain information on the internally managed portfolio of 13(f) securities for these plans. The forms contain the names, class, CUSIP, shares and value of each security. I managed to locate the 13F filings of 20 state-administered plans and used the filing closest to January 1, 2000 (a date near the midpoint of my sample) in my analysis. Using these filings, I matched each CUSIP to a company's headquarters using the COMPUSTAT database.

I compare the fraction of a state's portfolio invested in-state to the relevant in-state investments of the SP500 in Table 2. I find a small, statistically insignificant degree of overall bias, with an average of 0.31% of the portfolio's overinvested instate. No plan has more than 1.9% of the portfolio overinvested in state, and five out of the twenty portfolios are underinvested in-state. Similarly small results are found when comparing each plan's in-state investment to the investment made in that state by the other plans in my sample.

[15] A more detailed description of 13(f) filings can be found on the SEC website, at http://www.sec.gov/divisions/investment/13ffaq.htm

My newly constructed panel of investment returns enables another method of measuring the degree of in-state bias. Table 3 present the results of a series of regressions run using state retirement system returns as the dependent variable. Each of these regressions contains fiscal year fixed effects and the standard errors clustered by state.

For the first test, I construct a state industry-weighted return by multiplying each state-year's industry shares by the corresponding Fama-French industry portfolio.[16] Annual state industry shares are calculated from the BEA GSP series. Column (1) in Panel A indicates that this industry-weighted return has no additional predictive power for plan performance, with a point estimate close to zero. This is evidence against the hypothesis that systems overinvest in state specific industries.

In the second test, I construct value-weighted portfolios of the companies in the CRSP data set headquartered in each state. Though the number of companies in these portfolios varies considerably by state and year, the results are consistent with those found in column (1). Once again, the state-specific portfolio has no predictive power after controlling for time fixed-effects. This indicates that state plans do not significantly overweight companies headquartered in state.

Columns (3) and (4) run the same type of regression, with controls added for contemporaneous state income and employment growth. Though this may not be surprising if returns are forward looking, the lack of any meaningful relationship is reassuring. Again, there appears to be no appreciable relationship between idiosyncratic plan returns and contemporaneous state-specific economic shocks. Note that this is also a good test of the hypothesized channel, namely that pension windfalls affect income and employment only through future state spending. These windfalls have no relationship between contemporaneous changes in income, but are associated with increases in spending and income in the following state fiscal year.

In Column (5), I narrow the sample to plans with fiscal years ending in June. I then calculate the average return of all the systems in neighboring states. There are large regional differences in industry concentration. If state plans over-invest in state-specific industries, these 'neighbor returns' should have predictive power. Once again, though, my regressions yield small and statistically insignificant point estimates. Thus, my newly constructed data set on public retirement system investment returns demonstrates that in-state bias has little, if any, effect on the variation in returns.

Finally, I present a series of four graphs (Figure 4) exploring the relationship between state industrial share and portfolio correlation for four major industries. These figures show no significant relationship—either positive or negative—between concentration in these industries and the correlation of plan and industry

[16] Details on the aggregation of industries and my method for handling the SIC/ NAICS transition in 1997 are recorded in the online appendix. Data on the industry portfolios are available on Kenneth French's website, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

returns. This is further evidence that the variation in returns across plans is unrelated to the economic shocks within the administering state. Therefore there does not appear to be any a priori problem with using these returns as an instrument for state and local government spending.

3.1 Political Influence on Returns

Panel B of Table 3 addresses a secondary set of concerns, namely the possibility that political factors within a state might affect both its economic climate and its retirement system's returns. For example, corrupt politicians may cause both pension funds to earn low returns and the state to display poor economic growth. [17] Though a time-invariant corruption effect would be subsumed by the state fixed-effects included in later specifications, time-varying levels of corruption are a threat to the identification strategy.

I first test this hypothesis using the Glaeser-Saks (2004) corruption index, which measures the number of corruption convictions per capita in a state from 1990-2002. Though this variable does not vary over time, evidence on the level of corruption can be indicative of the importance of the time-varying component. While the coefficient is negative, the point estimate is small and insignificant. In column (2), I test whether or not returns are different for systems during gubernatorial election years. The hypothesis is that the political cycle may impart time-varying incentives for politicians to interfere with government retirement system. Once again, the measured effect is small and insignificant. Similarly small and insignificant effects are demonstrated for the governor's party in column (4) and the presence of an independent investment council in column (5).[18] Column (3) uses political contributions data collected from the National Institute on Money in State Politics.[19] The instrument is the share of political donations contributed by public employees' unions averaged over all the state elections for which there is data. This variable, which again does not vary over time, is designed to measure differences in the political power of these unions across states. The coefficient in the table indicates that the political influence of public sector unions has little impact on the returns earned by these plans.

---

[17]One notable example of this involves the Ohio Worker's Compensation Fund (not a retirement fund) between 1998 and 2005. In March of 1998, an Ohio GOP activist and donor was awarded a $50 million dollar contract to invest in rare coins on behalf of the fund. Many of these coins were 'lost', and the state ultimately recovered only $13 million of the original investment.

[18]In unreported tests, I similarly examine the effect of other years in the gubernatorial election cycle and a variety of party control measures. In none of these specifications were the estimated coefficients economically large or statistically significant.

[19]http://www.followthemoney.org/.

4 Building the Instrument: Excess Returns

Having established that the variation in returns across plans is not driven by in-state bias or political considerations, I now discuss how I use this variation to construct an instrument for state-level spending. The identification strategy in this paper relies upon isolating a component of pension plan returns that is exogenous to economic shocks. The absolute return earned by a system clearly fails to satisfy this requirement. One would not want to identify off the fact that all plans performed poorly in 2008, since this aggregate poor performance reflects aggregate economic shocks.

Simply demeaning the returns, however, is not sufficient to recover random variation.[20] There are considerable differences across plans in the choice of broad asset allocation. A forthcoming companion paper (Shoag 2010) demonstrates that these differences vary along a number of endogenous dimensions, including the stringency of state balanced budget restrictions and the level of employment and income relative to the state's mean. Endogeneity in the asset allocation decision could bias the second stage results, if for example, more cyclical states invested more heavily in high-beta assets. A number of plans have off fiscal years, so that including the full array of asset-class and fiscal year interactions in a demeaning equation is very taxing on the data. Therefore the approach used in this paper is to evaluate a plan's return relative to a comparison 'benchmark' return. This measure is commonly used by the plans to gauge their own relative performance.

When comparing two returns—the actual return and a benchmark—it is possible to divide the difference into a component that is due to asset allocation and a component that is due to relative performance within asset classes.

$$r_t^i - \bar{r}_t = \left(\sum_a w_{a,t}^i r_{a,t}^i - \sum_a \bar{w}_a \bar{r}_{a,t} \right) = \underbrace{\sum_A \left(w_{a,t}^i - \bar{w}_{a,t} \right) \bar{r}_{a,t}}_{Variation\ from\ Allocation} + \underbrace{\sum_A \left(r_{a,t}^i - \bar{r}_{a,t} \right) w_{a,t}^i}_{Variation\, from\, Performance}$$

For the reasons mentioned above, it is unclear whether the variation in returns that stems from broad asset allocation is exogenous. My baseline instrument controls for this by isolating only the variation in performance within asset classes. For each plan year I construct a plan-specific benchmark return equal to the investment return the plan would have earned had it maintained the same asset allocation and invested solely in aggregate benchmarks. For example, for a plan invested in 50% equities and 50% fixed income, the benchmark return would be the average of the SP500 and the Barclays/Lehman Brothers Aggregate Bond

[20]In practice simple demeaning does return similar first and second stage results, with point estimates of \$0.17 and \$1.6, respectively. This procedure does, however, results in a significant loss of precision.

Index. The asset classes controlled for in the baseline version include equities, fixed income, cash, real estate and mortgages, and 'other.'[21]

I subtract this benchmark return from the actual return earned, and use only this difference in constructing the instrument. This 'excess' return is scaled by the initial size of the portfolio, measured in real dollars per capita in the state, to deliver a measure of excess funds. These excess funds are then summed across plans to deliver a measure at the state-year level.[22]

The justification for using these excess funds as an instrument stems from a weak form of market efficiency. The underlying assumption is that, after conditioning for allocation, one cannot expect to outperform the market. Given the small degree of in-state bias, it is also unlikely that any outperformance within asset classes is due to factors that would bias the second-stage estimation (i.e. economic conditions within the state). Therefore there are compelling reasons to expect this source of variation to meet the identifying assumptions.

In Figure 5, I present a graphical analysis of the gap between realized and benchmark returns for the baseline instrument. As is evident in the graph, the excess returns are approximately normally distributed and centered near zero. This pattern is consistent with the desired 'random' nature of the shocks. There is little persistence in this measure of excess returns. The estimated AR(1) coefficient, after controlling for year fixed-effects, is 0.007 with a standard error of .04. An F-test on plan-dummies returns a p-value of .98, decisively failing to reject the hypothesis that plans consistently over or underperform their benchmarks.[23] Additional evidence of the 'random' nature of these shocks is presented in Figure 6, which displays the quintiles of these shocks for three years near the middle of the sample. These graphs show that there is considerable geographic and time series variation in which states receive 'high' and 'low' treatments of this instrument.

Summary statistics on the final instrument are presented in Table 4. These summary statistics demonstrate that these shocks (with a de-trended standard deviation of \$149) are large relative to the standard deviation of de-trended spending (\$341) and de-trended income (\$1,077). These large shocks are the product of both the large size of these pension plans and the substantial idiosyncratic component of returns. The size

[21] The Benchmarks used are the SP500, Barclays/Lehman Brothers Aggregate Bond Index, 3-Month Treasury bill, NCREIF Property Index, and the Wilshire 5000 Index. In the online appendix, I present the results using alternative benchmarks. The choice of benchmarks has little influence on the results.

[22] System fiscal years do not always match state fiscal years. In the event of mis-match, excess funds are assigned to the unfinished state fiscal year. For example, a plan with a fiscal year ending in December in a state whose fiscal year ends in June would assign the windfall earned from 1/1995-12/31/1996 to the7/1/1996-6/30/1997 state fiscal year.

[23] Details on these tests are presented in the online appendix.

of these shocks is important given the natural variation in income and employment and the short window of available data.

4.1 Robustness across Alternate Constructions

The availability of the data played a large role in determining the coarseness of the baseline asset class controls. To show robustness across constructions, I present results using additional forms of the instrument. The first version refines the asset classes controlled for, splitting equity and fixed income in to foreign and domestic components and using a private equity/venture capital plan-year average to benchmark the category "other." This is only possible for a subset of years and plans. As will be demonstrated below, this refinement has little impact on the results. Data availability prevents me from constructing even finer asset class controls, but it is difficult to imagine endogenous factors influencing the relative weights on finer asset classes in a systematic way.[24]

In addition to refining the asset class controls, I show robustness by varying the sources of variation included in the instrument. I first construct two broader measures that, while maintaining the variation from performance, also include variation from differences in asset allocations. To capture this broader source of variation, I subtract from each plan's return an 'average' return or macro benchmark that is no longer plan specific. These macro benchmarks only remove macroeconomic shocks and preserve variation due to differences in allocation.

I construct two such macro benchmarks. Benchmark A consists of a 70%/30% Equity/Fixed Income portfolio[25] and is used to capture the largest possible amount of variation. Benchmark B consists of the yearly average weight across plans for the equities, fixed income, cash, real estate, and 'other' classes. This benchmark narrows the variation included to only the differences in allocation relative to other plans that year. Both benchmarks are constructed with the same set of national indices used above. Once again, these benchmarks are subtracted from the actual return earned, generating a new 'excess' return. This 'excess' return is once again scaled by the size of the portfolio, measured in dollars per capita in the state, to deliver a measure of excess funds.

I further demonstrate robustness using one last variant of the instrument. This version captures only the variation in allocation decisions across plans. To construct this instrument, I use each plan's annual weight on the equities, fixed income, cash, real estate and other categories and then multiply these weights by

[24]If allocation decisions are endogenous down to the level of specific assets themselves, there is no exogenous variation off which to identify movements in government spending. This concern is unlikely to be realistic given that it generally only the allocation across broad asset classes that is set by the system's board.

[25]This allocation is a simplified version of the Fidelity Freedom funds allocation for middle-aged workers, and is a representative simplification of the weighted allocation across public systems.

national benchmarks. As in the prior construction, this gives me a plan specific benchmark. I then subtract off the macro Benchmark A (70% Equities, 30% Fixed Income) to capture variation in the plan benchmark relative to the aggregate returns. This version of the instrument captures only the variation across plans caused by differing asset allocations. By doing so, it removes all concerns about bias due to over-investment in-state. Despite using a largely orthogonal source of variation,[26] I show in the next section this instrument has similar effects on state government spending. This is further evidence that the connection between 'excess returns' and state spending runs through the hypothesized channel rather than through spurious correlation.

4.2 Testing for Endogeneity within Asset Classes

Before proceeding, I also perform one additional validity test based on the confounding example mentioned above: namely the possibility that within asset classes, states whose economies are more correlated with the market overinvest in assets with high betas. The baseline instrument controls for differences in the weights assigned to equities across states, but endogenous investment patterns within the equity class could bias the second-stage results. Though I cannot test this hypothesis for all the plans in my data set, I examine this possibility using CUSIP level data for the plans that filed 13F forms. Testing this proposition requires some finesse in measuring the correlation between a state's economy and aggregate market returns. This correlation is a function of both the underlying economic properties of the state and the impact of the state's retirement system investments. In testing for this additional bias, I need to isolate only the covariance caused by underlying economy. I do so by using national industry betas at the 2 digit SIC level collected from Damodoran (1996) and state-SIC industry weights from the BEA. Figure 7 shows the relationship between a system's 13F portfolio beta and the industry-weighted beta of the administering state. As evidenced by the graph, the relationship is only weakly positive (and negative if one excludes the small Montana portfolio). It is never statistically significant.

[26]The broadest measure of idiosyncratic returns, which uses returns relative to the Macro Benchmark A, is the sum of the baseline performance-only measure and this measure which uses only the variation allocation. A variance decomposition of the broadest measure reveals that 39% of the variance is due to the variance in performance, 49% of the variance is due to the variance in allocation and 12% of the variance is due to their covariance. The correlation between these two components is .14.

5 The Effects of Excess Returns on State Government Spending

"State-subsidized pension funds for Pennsylvania state workers and teachers said last week that their 2007 investment returns were far in excess of U.S. stock and bond market indexes... SERS Chairman Nicholas Maiale said his system's numbers put it "among the top 5 percent of large public funds." ... The high returns enabled SERS to slow the expected increase in the state's payroll contribution" –Philadelphia Inquirer (5/2/2008)[27]

5.1 Spending Effects

In this section, I show that the excess funds earned by a state's retirement system affect the state government's spending. This is the 'first-stage' in the instrumental variables strategy, and I demonstrate the strength of this relationship through a number of robustness checks. I also demonstrate that the effect of excess fund shocks on state tax receipts is small and that there is little evidence of tax cuts. This section motivates the use of excess funds as an instrument for spending.

My standard regression specification is

$$g_{i,t} = \alpha_i + \gamma_t + \beta * Excess\ Funds_{t-1} + \varepsilon_{i,t}$$

where spending and excess funds are measured in levels and real dollars per capita. The unit of observation is a state-year, and the standard errors are clustered at the state level. This controls for arbitrary serial correlation at the state level. The state and year fixed effects are designed to remove baseline differences across states and common shocks. The regressions are weighted by population. Unless otherwise mentioned, all the following regressions include these controls and use this functional form.

The results for the baseline specifications, as well as for the additional instrument definitions, are presented below in Table 5. The baseline specification (column 1) returns a point estimate of .43, indicating that spending rises by 43 cents in the year following a pension windfall of one dollar. This robust relationship can be seen graphically, and is presented in Figure 8. Column (2) modifies this regression by expanding the data set through 1984, with little change in the results. Columns (3) and (4) use the broader versions of the instrument that measure returns relative to the 70-30 and time averaged allocation respectively. Once again the coefficients are large and in-line with the baseline estimate, though the inclusion of the questionably exogenous variation from allocation does reduce the point estimate. Column (5) uses the baseline instru-

[27] These figures are for the calendar year. The PA Public School Employees earned 22.9% in FY 2007 (relative to a benchmark 17.1%), and the PA State Employees earned 16.4% (relative to a benchmark 13.9%). Combined with a small loss for the PA Municipal Retirement Fund, the de-trended windfall for 2007 was $354. PA's spending in 2008 was $70 above trend implying a ratio of 20 cents per dollar of windfall gains.

ment with refined asset classes controls.[28] The idiosyncratic returns measured against this more detailed benchmark have similar and somewhat larger effects. Last, in column (6), I estimate the effect of windfalls that are derived solely due to the variance in asset allocation. This measure of pension windfall completely removes the variation in performance, and hence is only weakly correlated with the benchmark instrument (correlation =.14).[29] Despite using a different source of variation, column (6) demonstrates that these windfalls generate similar spending responses. The results are further evidence that the channel from windfalls to spending is real and that the identification strategy is valid.

It pays to be careful about the power of standard statistical tests when using a panel of state-year observations. To check the specificity of my baseline specification, I conducted a series of Monte Carlo placebo simulations. In these simulations, I randomly reassigned (1) the excess funds series across states and years, (2) the excess funds variable across years within a state, and (3) the time series of excess funds across states. In Figure 9, I plot the CDF of the placebo coefficients along with the true point estimate. In all cases the true point estimate lies well outside the range of placebo estimates, suggesting the tests are not under specific.[30]

A second concern when dealing with a panel of states is the issue of spatial correlation. A recent paper by Barrios, Diamond, Imbens and Kolesar (2010) indicates that this should not be a problem given the random assignment of treatment along with the use of clustered standard errors. I confirm their theoretical finding by calculating the Moran's I statistic for the baseline specification (1). The result, a point estimate of -.02 on a scale from -1 to +1, indicates little spatial correlation in the errors, and is not statistically significant.

The unpredictable nature of these excess returns implies an additional validity check, namely that future values of the excess funds series should not affect today's spending. Table 6 demonstrates this result using detrended spending and de-trended excess funds. As demonstrated in columns (1) and (2), future values of the detrended windfall series have little predictive power on in-period spending. Columns (3) and (4) demonstrate that similar tests using the lagged windfall value deliver the expected strong positive relationship found in the first stage.

Further robustness checks on the first-stage regressions are presented in Table 7. Column (1) exploits the cross sectional variation in excess funds to run a series of 21 year-by-year regressions of spending on lagged

[28] The new benchmarks used are the MSCI Global Index for Foreign Equity, the JP Morgan Global Bond Index for Foreign Fixed Income, and the Cambridge Associates PE and VC indices.

[29] The existence of a positive correlation implies that plans that performed better within asset classes also chose better allocations across classes. This may not be surprising if there is some talent in portfolio management.

[30] These simulations similarly confirm the specificity of the test. In the most conservative simulations, which only reassign the complete time-series, 2.3% of the simulations resulted in coefficients that were statistically significant and positive, in line with the desired specificity of the test.

excess funds. The estimates from these regressions are then used to construct a point estimate and standard error, as in FamaMacBeth (1973). The recovered coefficient is statistically significant and comparable to the pooled specifications. Column (2) reports the symmetric procedure across states. I run 48 state-by-state regressions, in which spending is regressed on a year trend and the lagged value of excess funds. The 48 coefficients are once averaged to create a point estimate and standard deviation in accordance with Fama-MacBeth. The relationship between excess funds and spending is once again significant and comparable in magnitude. These tests indicate that the first-stage relationship is present in both the cross-section and time-series, and robust along both dimensions.

Columns (3) and (4) present robustness results when including state trends and state quadratics. The first stage relationship remains strong in both cases. Column (5) uses a rescaled instrument to present comparable results when the dependent variable is in logs. Column (6) presents comparable results using the first difference of both the instrument and dependent variable. Last, column (7) presents the baseline specification with five additional controls: a state-level leading indicator created by the Philadelphia Federal Reserve, the annual per-capita dollars received from the Tobacco Master Settlement, the annual per-capita dollars received by the state in Federal grants, the initial balance in the state's general and 'rainy day' funds and the initial size of the state's measured pension assets. With the addition of these controls, the precision of the first stage is enhanced, with the first stage F-stat on excess funds increasing from 14 to 23.[31]

In Table 8, I further explore the relationship between state government spending and the excess funds earned by a state's pension fund. In columns (1) (2) and (3) I examine the effect of spending from a dollar of excess funds over time. I show that the largest impact on spending occurs in the next fiscal year and declines steadily over time. After three years there is no longer any significant affect on expenditures. Column (5) presents a robustness check in which I use only states in a current budget cycle. In column (4) I present the results which allow the effect of excess funds to differ based on whether or not they are greater or less than zero. The intuition is that states may be less likely to cut spending or increase contributions in the face of a negative shock than they are to increase spending and cut contributions after a positive one. This conclusion is indeed borne out in the data, as the coefficient on positive excess fund in column (4) is larger and more statistically significant.[32]

[31] Recent research by Jose Luis Montiel Olea and Caroline Pflueger (2010) indicate that the conventional cutoff for weak-instruments (F>10) may be inappropriate in the presence of heteroskedasticity and autocorrelation. A first-stage F-statistic of 23 is sufficient to rule out 10% worst-case bias even under their more restrictive critical values.

[32] An F-test does not, however, reject the equality of coefficients. Thus the political economy interpretation should be taken with a grain of salt.

The year-ahead coefficients are larger than expected, given the various actuarial safeguards designed to insulate state budgets from large return shocks. This indicates that states find a way to realize these gains despite these provisions—either by explicitly contravening them[33] or by spending in anticipation of lower contributions. I explore this issue in depth in Appendix A.

5.2 Levels of Spending

The next series of results explores the impact of excess pension fund returns on the different levels of government spending. The results indicate that the bulk of the spending response occurs at the state level, with Federal spending unaffected by the shocks and local spending responding primarily through changes in state-level grants.

These facts are demonstrated in Table 9. Columns (1) and (2) compare the impact of excess funds on total state and state and local government spending (net of intergovernmental transfers). The point estimates are broadly similar across regressions, indicating that much of the reaction to these returns occurs at the state level. Column (3) examines the effect of excess funds for local government spending alone. The coefficient is of sizeable magnitude, but is no longer significant. Column (4) instruments for the value of grants given by state governments to local governments and finds a large effect roughly comparable to the overall increase in local government spending. Column (5) explores the effect of the excess funds shocks on local government spending after controlling for state grants. The estimates are more than three times smaller than the state-level results, and are not statistically distinguishable from zero. Though local level spending appears unrelated to the excess funds series after controlling for state-level spending, I demonstrate below that including this spending in my endogenous variable does not affect my second-stage regressions. Columns (6) and (7) explore the impact on federal grants to states and direct federal spending respectively. Both regressions indicate that these types of spending are unaffected by these shocks. Column (8) demonstrates that the association between excess funds and state level spending is unaffected when controlling for federal grants to the state.

[33] This does happen, as there are frequent revisions to the laws governing state pension accounting. One egregious example occurred in New Jersey in 2001, when the legislature decided to value pension assets at their value on June 30,1999 irrespective of both the actuarial and market values. The legislature then used the 'windfall' to fund a benefit and salary expansion. (Mary Walsh,"N.J. Pension Fund Endangered by Diverted Billions", NY TIMES 4/4/07). Other identifiable examples of accelerated realizations include CALPERS in 1991, New York in 1989, and an attempt in Ohio in 2009.

5.3 Composition of Spending

The random nature of excess funds shock allows me to explore what parts of a state's budget absorb the financing pressure. I run a series of seemingly unrelated regressions on the components of spending. Table 10 presents these estimates at the state level, broken down by functional categories of expenditure. Unreported results using total state and local spending produce nearly identical results.

As demonstrated in the table, the largest response comes from spending on education. It is tempting to attribute this to the fact that a large proportion of the funds in state-administered plans are held on behalf of teachers. The composition of spending generated by the windfall shock, though, is actually quite similar to the average state-spending composition. To explore if there are differences in spending across plans, I reweight each system's excess funds by the proportion of members who are educational employees (as reported in the PENDAT surveys). Using this version of my instrument, I find a slightly larger estimated effect on education (0.16^{**}, SE=.06) An F-test cannot, however, reject the equality of these coefficients.[34]

At an operational level the largest response occurs in the current expenditure category, which accounts for 39¢(SE=.11) out of a total spending response of 43¢. The impact of the excess funds shock on capital expenditure, interest payments and subsidies is much smaller and not statistically significant. Direct government salaries go up by 8¢ per dollar of excess funds. This figure is not estimated precisely, but is still significant at the 10% level.[35]

5.4 The Absence of a Tax-Rate Response

Though an excess funds shock requires an adjustment in a state's budget at some horizon, this adjustment does not need to come solely through variations in spending. It is possible that state governments cut their tax rates after positive shocks and raise them after shortfalls. Using the same specification as before, I explore the effect of excess fund shocks on taxes.

Columns (1) and (2) of the first panel in Table 11 show the effect of $1 windfall on general revenue and personal income tax revenue is positive and significant. Rather than reductions in revenue due to falling tax rates, I find that revenues actually increase. Though this result may seem puzzling, it corresponds nicely with our later finding of large and significant multipliers on government spending. Revenue actually rises after an excess funds shock due to the extra income generated by increased government spending. Using my baseline multiplier estimate of 2.1, combined with the baseline first-stage estimates and a calibrated state

[34] If the windfalls from different types of retirement systems do display systematic differences, then it would be possible to separately identify the income and employment effects of different types of spending. Work on this extension is ongoing.

[35] Details on these specifications are in the online appendix.

income tax of 7.5%, I expect to find an increase in personal income tax revenue of approximately \$.07. This is well within the confidence interval of the actual estimates in column (2). This explanation corresponds nicely with the finding in column (3), which shows no response in the state tax burden (total taxes paid relative to income), despite the increase in revenue.

It should be noted that this result, that tax revenue (and presumably income) rises in the years following positive idiosyncratic state pension returns was already noted by Brown, Pollet and Weisbenner (2009). This paper demonstrates a statistically significant correlation between detrended tax revenue and lagged returns, but not between tax revenue and contemporaneous returns. Brown et al. interpret this as indirect evidence of in-state bias and a forward-looking component of returns. Given the demonstrably small levels of in-state bias, this paper suggests an alternate interpretation of their results—namely that income rises due to a multiplier on state spending. Interpreted in this light, their finding could be interpreted as independent evidence supporting this paper's results.[36]

When assessing the ability of pension returns to isolate spending rather than tax shocks, the more interesting hypothesis is that state tax rates (rather than revenue) will fall in response to an income shock. I explore this in two ways. Column (4) in Panel A and Column (1) in Panel B regress the state sales tax rates and the marginal property tax rate on the excess funds shock per capita. These regressions, as well as unreported specifications looking at future values of these rates, find no relationship between excess funds shock and tax rate adjustments. State income tax rates have many institutional details and do not lend themselves easily to regression analysis. Thankfully the National Association of State Budget Officers (NASBO) conducts an annual review of all revenue changes enacted at the state level. This review, in the fall *Fiscal Survey of the States* assigns a fiscal year expected change in revenue associated with each change in policy. I use these legislated changes, scaled to the real per capita level, to investigate whether excess fund shocks affect the effective level of personal income taxes. Once again, I find no effect on personal income tax rates and legislated total revenue either contemporaneously or in the future from excess fund shocks.

The lack of a measurable effect on tax rates motivates using these pension windfalls as an instrument for spending alone. The apparently small effect on taxes makes it unlikely that the response I measure is caused

[36]To assess whether it is plausible that this correlation (measured at .29 in Brown, Pollet and Weisbenner after state and year fixed effects) could be driven by in-state bias, note that the partial correlation of the portfolio and in-state income can be written as:

$$\rho^{Portfolio,\ Income} \approx fraction_{in-state} * \rho^{In-State,\ Income} \frac{\sigma_{In\ State}}{\sigma_{All}}$$

Calibrating to the CRSP HQ- portfolios, I find that the ratio ($\rho^{In-State, Income} \frac{\sigma_{In\ State}}{\sigma_{All}}$) is approximately equal to .34.

This implies in-state bias on the order of 86% of the portfolio. Even assuming the in-state portfolio is perfectly correlated with in-state income and that in-state investments have a standard deviation of 3 times the annual SP 500 standard deviation (14.8% in my sample), this would still imply bias 9.7%. These estimates are dramatically higher than the level of bias found in the data.

by contemporaneous tax reductions. There is, however, a concern that the windfall itself may have a direct effect through the anticipation of future tax rate changes. If agents other than the state government make decisions based on the funding status of state pension plans, then the estimates of the effect of spending on income and employment may be biased. In practice this issue is likely to be second order,[37] and I attempt to control for the direct effects below. These tests yield very similar 'pure' spending multipliers (i.e. holding the size of windfall constant), confirming the conjecture that the direct effects are small.

6 Using Windfall Returns as an Instrument

6.1 Income Effects

Having established that idiosyncratic returns on state pension funds are strong predictors of future state spending, I now use these returns to instrument for spending in a regression on state level personal income. The standard specification is once again in dollars per capita, weighted by population, with controls for state and year fixed effects.

The baseline results are presented in the first column of Table 12. The point estimate (2.12), which is significant at the 5% level, can be read as dollars of income per dollar of spending. In other words, each dollar of exogenously assigned windfall spending generates $2.12 of personal income.

The remainder of Table 12 contains robustness checks on the baseline specification. Weak instruments are a common source of bias in instrumental variables estimation. Though the F-statistic on the first-stage instrument in the baseline specification (13.9) lies above the standard threshold for concern (Staiger and Stock (1997) introduce a rule-of-thumb cutoff of 10), in column (2) I explore the robustness of the second stage to more powerful first-stage specifications. Specifically, I now including the battery of controls (Federal Reserve Leading Indicator, Rainy Day Fund Balance, Initial Pension Fund Size, and Tobacco Revenue) explored in section 5. Using these controls, the first-stage F-statistic is now 23, and there is little change in the point estimate. Columns (3) and (4) repeat the baseline specification, now using the macro Benchmark A and macro Benchmark B versions of the instrument. Both versions broaden the sources of variation used to estimate the first stage, including the variation in allocation relative to the unconditional benchmark and to the time average benchmark respectively. Both instruments once again deliver point estimates similar to the benchmark regression (2.08 and 1.65), with the macro Benchmark A coefficient significant at the 5% level. The loss of precision under the macro Benchmark B construction stems mainly from a few outlying observations. To demonstrate this, I rerun the regression applying a filter that drops the largest 5% of the

[37] Collecting data on state pension fund returns required considerable time and effort for the researcher, making it unlikely that these relative excess returns are widely known by private agents.

detrended excess funds series. This regression, presented in column (5) once again displays a larger coefficient (2.47), this time significant at the 10% level.

The standard errors estimated in this table are clustered by state. In unreported regressions, these standard errors were estimated when clustering by year, by state and year (see Thompson 2009) and by year with state-specific AR (1) serial correlation correction. The standard errors were largest when clustering by state, and hence the significance levels reported here are conservative.

Panel B demonstrates the robustness of the second stage results across changes in specification. Column (1) displays the baseline regression after applying a filter that removes the largest 10% of the de-trended excess funds series.[38] Column (2) displays the regression with the inclusion of state-specific time trends. In both regressions the second stage coefficient of spending on income is close to the baseline (2.39 and 2.37) and significant at the 5% level. In Column (3), I estimate the equation with spending and income in logs, with the instrument scaled by the two year lag of state spending.[39] Though the precision of the first stage is weakened, the estimated elasticity is .38, which when interpreted at the mean of the two series, implies 2.86 dollars of additional income per dollar of spending. Column (4) presents the results with the instrument, spending and income all in first differences.[40] Again, though the first stage is weakened by the transformation, the results (2.05) are very similar to the baseline specification. Column (5) uses only the positive observations of the excess funds instrument. Despite limiting the source of variation, the second-stage point estimate (2.39) is quite close to the baseline and remains significant at the 10% level. The point estimates are reassuringly similar and stable across fairly major changes in specification.

The reduced-form relationship between income and excess funds can be seen visually and is presented in Figure 10. The axes in this figure are the detrended excess funds and detrended spending series, with each point representing the population-weighted mean value of four percent of the sample.

The second-stage results are not driven by any single year or state. Figure 11 plots a histogram of estimated second-stage coefficients when dropping one year (Panel A) and one state (Panel B). The range

[38] There are a number of reasons to suspect measurement error in extreme observations (e.g. incorrect reporting of calendar vs. fiscal year, off marking to market, etc.). As demonstrated in the table, this filter does not meaningfully impact the results.

[39] The two-year lag is used to remove any contemporaneous causality between the lag of excess funds and the lag of spending. I find little correlation in practice, and the results are identical when using the first lag.

[40] The dependent variable in the regression presented in Table 12, is income_t - income_{t-1} .As a placebo test, I run the same regression on income_{t-1} -$\text{income}_{t-2.}$. If the empirical approach is truly picking up changes in income associated with the change in spending, then the coefficient on the lagged change in income should be close to zero. The regression generates a point estimate on instrumented spending of -0.97 with a standard error of (1.5), reassuringly failing to reject the null of no effect at even the 10% level. This stands in contrast to the significant result found for the contemporaneous first difference, as demonstrated in Table 12. Details on this test are presented in the online appendix.

of estimates is reassuringly small, with the largest impacts stemming from dropping California (a reduction in the baseline second stage to 1.53) and Ohio (an increase in the baseline second stage to 2.77).

6.2 Instrumenting for State and Local Spending

As discussed earlier, the retirement systems tracked in the data are managed on behalf of both state and local government employees and receive contributions from governments at both levels. Though I showed in a previous section that significant spending response could be seen at the local level, failing to account for an increase in local government spending could in principle overestimate the impact of spending on income.

To explore this issue, I rerun the baseline and macro Benchmark A specifications instrumenting for state and local government spending in Table 13. In both baseline specifications (Columns (1) & (4)), the point estimate is large, significant and of comparable magnitude to the earlier baseline. The precision of the first stage estimate, however, is considerably weaker (F-Statistics of 4.9 and 4.7) and prone to weak instrument bias. I numerically calculate the weak-instrument robust Andersen-Rubin 90%-confidence intervals for both regressions. The confidence interval excludes zero in the first regression, but not when using macro Benchmark A instrument.

The weak-instrument problem for state and local spending is ameliorated through the use of a 5% filter on the excess funds series. After dropping the largest 5% of detrended excess funds series in absolute value, I once again recover the baseline estimate on income (2.28 and 2.24) significant at the 5% level with more power first-stage results (F-Statistics of 9.43 & 8.7). The fact that these estimates are so similar to the estimates obtained when using only state government spending confirms, as shown in an earlier section, that local government spending has a limited response to these shocks. Results identical to the baseline are also recovered by including the five controls used in Table 13 (as demonstrated in Column (2)), even without applying the trim.

6.3 Components of Income

Ideally one would like to be able to determine which components of income increase in response to these exogenous and unexpected spending shocks. Unfortunately data on consumption, investment, imports and exports are not available at the state level.[41] In Table 14, I offer suggestive evidence by exploring the impact of spending on proxies for consumption and investment. Column (1) demonstrates the effect of spending on retail sales, using data from the US Statistical Abstract measured for the calendar (not fiscal) year. It shows

[41] The Census does have foreign import and export data at the state level, but these series begin in 2006 and 2008 respectively. These short horizons make these series unusable for this exercise.

that a \$1 increase in spending raises retail sales by \$1.62. This effect is not precisely estimated, perhaps due to the calendar-fiscal year mismatch, but suggests a strong consumption response.

These positive findings for the consumption proxies stand in contrast to the results for investment. Column (2) uses the state-level manufacturing investment data set assembled by Wilson and Chirinko (2009). These data are derived from the Survey of Manufacturers and as before are measured on the calendar year, introducing measurement error. Still, the negligible coefficient suggests that additional spending has a small effect on in-state investment. This intuition is borne out in a second test, displayed in Column (3), which uses data on investment for the COMPUSTAT sample of firms. In this regression, I assign firms based on the location of their headquarters and construct an investment measure equal to their capital expenditure scaled by their previous year's book value of assets.[42] I then regress this ratio on the set of year and state dummies, and state level spending instrumented with the windfall variable. As in Column (2), the recovered effect on investment is small. The estimate in Column (3) suggests that for every \$100,000 dollars in per capita spending, in-state firms reduced capital expenditure by 2.3% of assets. This result qualitatively confirms the findings in Coval, Cohen and Malloy (2010) that the investment behavior of companies headquartered in state responds negatively to in-state government spending. Taken together these findings suggest that the bulk of the increase in income stems from increased consumption and not an increase in investment.

6.4 Employment Effects

Having demonstrated a strong and robust effect of spending on income, I now turn to employment. Data on in-state employment is collected from the BLS Current Employment Statistics (CES), a monthly establishment-based survey. Robustness results using data from the Quarterly Census of Employment and Wages (QCEW)[43] and from the household-based Local Area Unemployment Statistics (LAU).

The dependent variable is measured as employment per capita, and to enhance readability, the spending variable is now reported in units of \$100,000. The basic specification is as before, with spending and excess funds measured in 2009 dollars, state-and-year fixed-effects and standard errors clustered by state. Thus the coefficients in Table 15 can be read as the number of jobs generated per \$100,000 of spending. The inverse (dollar-per-job) is reported below.

As demonstrated in Table 15, spending has a robust positive impact on employment with the baseline estimates indicating that \$100,000 of spending generates an additional 2.89 jobs. Switching from the CES

[42] Firms report data over a range of fiscal year, many of which do not correspond to the state's fiscal year. As before, this timing problem is an additional source of bias.

[43] While more comprehensive generally, the QCEW fails to cover 900,000 state and local government employees. The CES tracks employment for these workers using outside sources.

series to the QCEW or LAU series changes the coefficient slightly to 2.84 and 2.58, as shown in Columns (3) and (4). The point estimate is robust across versions of the instrument (as demonstrated in the first column of Panel B) and across controls (Columns 2 and 3 in Panel B). As before, including local spending in the endogenous variable weakens the precision of the first stage. To account for this, I calculate the weak-instrument robust AR 90%-confidence interval ([1, 11.7] jobs per $100K) for this specification. Last in column (5) I present results which include an additional control for predicted state employment. This control is built following the methods developed in Bartik (1991) and Blanchard and Katz (1992)[44] Though the point estimates shrink a bit with the added control, there remains a large and significant employment effect.

The reduced-form relationship between employment and excess funds can be seen graphically, and detrended employment is plotted against detrended excess funds in Figure 12.

Using data from the LAU, I am able to decompose these employment effects into changes in labor force participation and changes in the unemployment rate. Once again maintaining the same basic specification (expressing the dependent variables in per capita terms), I show the bulk of the increase in employment stems from increased labor force participation. The point estimates in Panel A of Table 16 indicate that $100,000 in spending creates 2.85 new jobs, drawing 2.51 people in to the labor force and removing .33 people from unemployment. This increase in employment occurs both in the private and public sectors. The estimates in Panel B indicate that $100,000 of spending creates 2.15 private and .45 public sector jobs.

Using data from the BLS's Business Employment Dynamics survey, I can also decompose the increase in employment into the effect on gross job gains and gross job losses. The Business Employment Dynamics data measure job gains and losses using the administrative records of the QCEW. A gross job gain occurs when an establishment increases the number of employees on payroll relative to their report the prior month. A gross job loss occurs when an establishment reduces the number of employees on payroll relative to the previous month. In Panel C, I explore the effect of spending on net job creation, the difference between the two series. The point estimate indicates that an additional $100,000 of spending generates 3.69 new jobs. This point estimate differs slightly from earlier estimates due to the different time period and data source. Columns (1) and (2) decompose these net job gains into gross job gains and gross job losses.

As evidenced by the table, the bulk of the net employment increase stems from an increase in gross job gains. The point estimates in Column (1) indicate that an additional $100,000 of spending yields an increase of 6.56 job gains. The response of gross job losses is smaller and it too is positive. The point estimates

[44] For each state s, I calculate the change in employment for each three digit NAICS industry over the 49 states which exclude state s. I then weight these employment changes by the initial level of employment in the three digit NAICS industry in state s.

indicate that \$100,000 of additional spending increases the number of job losses by 2.87. It is important to note that these losses do not correspond to firings, but rather to positions removed or left vacant. The mild increase in job losses would be consistent with a fraction of the new job gains being filled by job-switchers. Thus the net employment estimates belie an even larger amount of churning, with some positions crowded out as a result of higher spending.

6.5 Persistence

Ideally one would like to be able to measure the impulse response of a shock to spending on income and employment. Unfortunately the response of the time-series to a one-time increase in spending is not properly identified in the data. As shown in Table 8, pension windfalls affect spending at multiple horizons. Without additional instruments or assumptions, the data can then be used only to identify the effect of one time series of shocks on another.

It is possible, though, to present suggestive evidence on the persistence of the income and employment effects. This evidence indicates that the persistence is small, and that the gains largely disappear when the spending stops.

Table 17, presents a cumulative measure of the income and employment effects. The dependent variable is the sum of income and employment over the relevant horizon and the endogenous explanatory variable is the sum of spending. If spending had long-lasting effects on the dependent variables, then the reported coefficients should be rising over time as the delayed effects cumulate and spending ceases to rise. The relative stability of the reported coefficients across years indicates that this is not the case, though there appears to be more persistence in employment than in income

Additional evidence that the persistence is small can be garnered by making assumptions on the parameters governing the impulse response. To examine the persistence of the income effects, I impose two assumptions: that income does not respond to future spending and that the effect of contemporary spending is constant across periods. This allows me to identify the next period's persistence parameter μ^1

Impulse Response with Imposed Assumptions:

	T=0	T=1
Spending	G	G_1
Income/ Employment Response	μ G	$\mu\ G_1 + \mu^1\ G$

Imposing these assumptions on the system estimation, I find that the effect of \$1 of spending in year zero raises the following year's income is \$0.46 as opposed to an in-year effect of \$2.12. Similarly, \$100,000 of spending in year zero raises employment in year one by 0.82 jobs, compared with an in-year increase of 2.89 jobs. In both cases, the following year effect is not statistically significant.

Under the assumptions imposed on the data, the stimulative effects of government spending are not persistent. While the data do not permit me to rule out long-lasting effects, the evidence available suggests that these effects, at least for income, are small at best.

7 When is the Multiplier Large?

7.1 Labor Force 'Slack'

Traditional Keynesian theories that generate large multipliers often assume that there is economic slack, in the sense that some factors of production (e.g. labor) are unproductively idle. These theories suggest that the impact of government spending on income and employment should be larger when an economy is functioning below capacity. To explore this mechanism, I divide my sample along a number of dimensions designed to measure the degree of slack in production. I then rerun the baseline second-stage specification on each half of the sample.

I consider two measures of labor force slack. Columns (1A) and (1B) in Table 18 explore the effect of spending when employment the previous year was low (high) relative to the state's mean. Columns (2A) and (2B) report the effects when labor force participation fell (rose) the previous year. Both sets of regressions indicate that spending has much larger income effects following periods of falling or low economic activity. The coefficients in columns (1A) and (2A) are much larger and more precisely estimated than the analogous results in columns (1B) and (2B). Baseline estimates indicate that an additional \$1 of spending in the face of economic slack generates \$3.00-\$3.53 of income. The comparable effect when factor slack is low is only \$1.61 to \$1.40 per dollar of spending, and these point estimates are not statistically significant.

Columns (1A), (1B) (2A) and (2B) explore the effect of economic slack relative to the state's baseline or previous years. In columns (3A), (3B), (4A), and (4B), I use divisions that measure the absolute rather than relative level of factor slack. Columns (3A) and (3B) split the sample based on whether or not employment exceeded 45.5%, and columns (4A) and (4B) split the sample based on whether or not labor force participation exceeds 50%. Whereas the previous regressions largely looked across years within a state, these divisions capture mostly variation across states. Nevertheless, the estimated effect of spending on income is similar across divisions. In both instances, the income effects are larger in economies with more economic slack. The results are somewhat smaller than in the first panel (2.6 and 2.3 vs. 3.5 and 3.0) but display a similar pattern. Interestingly, the first-stage coefficients are also larger in the "A" series of regressions. This indicates that pension windfalls are more likely to be spent in poor economic times. Similar results are also found when using other variables, like income, to classify times of economic 'slack.' Thus the data do seem consistent with the traditional Keynesian mechanism in this respect.

7.2 Tradables vs. Non-Tradable Industries

Many papers assume that government spending is concentrated in non-traded industries like services. The direct and indirect effects of government purchases are likely to be concentrated in industries whose production cannot be filled out of state.

Following Rodrik (2007), I use a broad division of industries that classifies agriculture, mining, manufacturing, and wholesale trade as tradable. Construction, utilities, transportation, retail, financial, real estate and other services and government expenditure are classified as non-traded. Using this division and the BEA regional income data, I show in Table 19 that the effects of government spending are heavily concentrated in the non-traded sector. Using the QCEW to divide employment along the same lines (though excluding direct government employment), I similarly find that employment and average wages also go up in the non-traded sector. Columns (4) and (5) indicates that each dollar of government spending raises average wages in the non-traded sector by $2.98, and that $100,000 of government spending adds 1.71 jobs in the non-traded sector. Columns (2) and (3) show the analogous effects in the traded sector are small and statistically indistinguishable from zero.

Using these results, I can decompose the contribution to overall income implied by the changes in employment and wages across sectors. At the mean level of employment in the non-traded sector, I find that the increase in wages in that sector accounts for $.86 of the increase in income. Similarly, updating at the mean wage in that sector, I find that the increase in employment in that sector generates an additional $.60 of income. The direct income effect of government spending is estimated at $.33 per dollar, which accounts for a total of $1.79 of total income.

7.3 Spillovers Across State Borders

While most studies look at multipliers at the national level, this paper explores state-level effects. As discussed previously, there are a number of reasons to anticipate different effects at the sub-national level. Chief among these is that there is likely to be some economic spillover, as the increased demand at the local level affects demand for the goods produced in neighboring states. These spillovers could induce either positive or negative responses in neighboring areas, depending on whether the increase in spending stimulates neighboring economies or draws away demand or resources.

I explore this issue by first measuring the income effect of in-state spending on neighboring out-of-state counties. To identify counties with strong economic ties across states, I use the Census's 1990 Journey to Work dataset. I divide counties in the neighboring states into two groups, based on whether they have greater than 10% of their workforce commuting to the treated states. I then compare the effect of in-state spending on the two groups.

As demonstrated in columns (1) and (2) of Table 20, state spending in state has a strong stimulative effect on economically linked counties across the border. The measured income response in those counties is similar to the income response of counties within the state borders. These income effects are not present in the counties with weaker economic ties. This suggests that the spillover on near-neighbors is positive, and it is reassuring that the in-state income effects are not caused by a quirk in the state-level income data.

Column (4) of Table 20 explores the possibility that state spending might generate migration across state borders. Using data from the IRS on changes in tax return filings, I find little evidence of people moving in response to pension windfall shocks. The point estimate indicates that an additional $100,000 of spending causes .5 people to move into the state and is not significant. This small response is not surprising. Though the windfall shocks are large relative to the annual de-trended standard deviation of income, they are temporary and small relative to the costs of moving.

Lastly column (5) of Table 20 explores the extent to which spending shocks affect in-state housing prices. Keynes (1929) and Ohlin (1929) famously debated the extent to which cross-state transfer would result in the appreciation of the recipients' terms of trade. Lane and Milesi-Ferretti (2000) and others have shown that this debate can be mapped to the question of whether transfers from one country to another would result in an appreciation in the price of the recipients non-tradable goods. Column (5) tests whether or not the increase in government spending is indeed capitalized as an increase in housing prices. The specification uses the first difference of the windfall and spending to look at the percent change in the Federal Housing Finance Agency housing price index. The reported coefficient indicates that an additional $1 of spending per capita raises housing prices by 0.01%. Thus there is some evidence that transfers do affect prices.[45]

7.4 The Windfall or Government Spending: Suggestive Evidence

One concern when evaluating these estimates is that it is unclear whether the income effects are the product of the state spending or the windfall returns themselves. The IV estimates above implicitly assume that, save for the impact on spending, the windfall return would have no effect on in-state income. While introspection may suggest that the direct effect of pension windfalls is small, it is nevertheless impossible to test this conjecture directly without a second instrument. Despite being unable to test the direct effect in a concrete manner, I can provide suggestive evidence that the direct effect is unimportant.

[45] I am in the process of building detailed state-level price series, which I will use in future research to explore this issue in more depth.

The first such test comes from comparing the reduced form income effect of windfalls across states with large and small spending responses. As in the Fama-MacBeth regression in Table 7, I run state-by-state regressions of spending on a year trend and the lagged excess funds series. This procedure results in a state-by-state estimate of the spending response to windfalls, with a median spending response of 15.6¢ per dollar. I then run state-by-state regressions of the reduced form second stage, regressing income on the lagged excess funds variable and a year trend. This results in a state-by-state estimate of the income effect of windfall shocks.

In Table 21, I present the results of a regression of the reduced-form second-stage coefficients on the first-stage coefficients. The reported coefficient measures the extent to which increasing the amount of the windfall being spent enhances the effect of the windfall on income. In other words, it is a measure of how much the spending itself matters holding the size of the windfall constant. The coefficient displayed in the table is nearly identical to the baseline second-stage result. This suggests that the large income effects found in section six are driven entirely by spending and not the windfalls themselves.

The spending response measured in the above example is certainly endogenous, and the multiplier 'holding the windfall constant' is not cleanly identified here. Still, these results are suggestive that the spending effect dominates the direct one.

This result is further supported by an additional test, using the stringency of state-balanced budget requirements to isolate 'exogenous' spending responses. All states except for Vermont operate under a balanced budget requirement, but these rules vary considerably in severity. The Advisory Commission on Intergovernmental Relations (ACIR) rated states' requirements on a scale of 1-10, from least to most severe in a 1987 report. Beginning with Poterba (1994), a number of studies have used this rating to examine the budgetary impact of these rules. Generally these studies conclude that these rules have a large impact on a state's spending response to shocks.

This difference is manifested in the way states respond to pension windfalls as well. Following Clemens and Miran (2010), I divide states in to strong (ACIR index>7) and weak rule groups using the ACIR index. I then include an interaction term for weak-ruled states in my baseline regression of spending on year dummies, state dummies and the excess funds series.[46] The intuition is that weak ruled states may be better able to accelerate spending from windfall gains. This intuition is confirmed in Table 22, as the interaction term returns a coefficient of \$0.43 and is significant at the 1% level. Thus weak-rule states spend, on average, forty cents more than their strong-ruled counterparts.

[46]The dummy for strong-ruled states is subsumed by the state dummies.

I then estimate an analogous version of the reduced form equation for income, identical to the earlier specification save for the introduction of a weak-rule lagged excess funds interaction term. Once again, I hope to recover a multiplier estimate 'holding the size of the windfall constant' by comparing the magnitude of these two terms.

In the modified reduced form equation, the interaction term returns a coefficient of $1.1. Comparing the ratio of the coefficients yields an estimate of the multiplier, holding the size of the windfall constant, equal to 2.75 (1.1/.4). This ratio just fails to obtain statistical significance in the system estimation. Nevertheless, its similarity to the earlier results suggests a small direct windfall effect. Once again the available evidence indicates that the large multipliers found in section six are driven by state spending and not directly by the windfalls themselves.

8 Comparison to the Predictions of a Standard Macro-Model

How do the empirical findings presented above compare with the predictions of a standard macroeconomic model? This section shows that the standard frictionless model does not match the central result of a multiplier larger than one, even in the case of 'windfall' financing. The failure of the model stems from the standard intratemporal optimality condition between labor and consumption. I then explore the effect of modifying this condition by incorporating either sticky prices or sticky wages. I find that adding these frictions allow the model to match the main empirical findings of a multiplier greater than one and an increase in employment.

To understand why these frictions change the predictions of the model, consider (as in Gali, Lópes-Salido and Vallés 2005), the standard intratemporal optimality condition:

$$mpl_t = mrs_t = (\alpha - 1)\, h_t = \sigma c_t + \phi h_t$$

where $\sigma,\ \phi > 0$ measure the curvature of utility with respect to consumption and labor and $\alpha \in [0, 1]$ measures the decreasing returns to labor. This equation illustrates the standard 'wealth effect' on labor supply: when consumption increases, the labor supply curve shifts in. When this equation holds, it is impossible for both consumption and output to rise. Thus government spending shocks cannot generate the large multiplier found in the data.

Other papers attempting to deliver large spending multipliers use various modeling devices (e.g., sticky prices, rule-of-thumb consumers, GHH preferences, etc) to alter this intratemporal optimality condition. In this paper, I modify this condition by assuming price and then wage markups. With these assumptions, the labor-leisure tradeoff becomes:

$$(\alpha - 1)\, h_t = \sigma c_t + \phi h_t + \mu_t^p + \mu_t^w$$

where μ_t^p and μ_t^p are the log price and wage markups. I show that when prices and wages are sticky (pre-set), endogenous movements in these markups allow the model to generate a multiplier greater than one.

This section proceeds as follows. I first describe the way I model the empirical experiment. I then examine the effects of this 'windfall spending experiment' in five models. In Model (1) I consider the windfall experiment in an endowment economy. This clarifies the accounting differences between personal income (the variable used in the empirical work) and output. Model (2) considers the same experiment in a standard closed economy with endogenous labor supply. For the reasons outlined above, this model is unable to generate a multiplier greater than one and an increase in hours worked. In Model (3), I generalize Model (2) to an open economy with trade. This is a natural extension when modeling a shock at the state level and introduces potentially important terms of trade effects. I show that this generalization does not change the basic responses found in Model (2). Lastly, in Model (4) and (5) I incorporate sticky (pre-set) prices and wages, respectively, in to Model (3). I show that with these frictions the models are able to deliver an increase in employment and a multiplier above one.

As discussed above, the spending shocks identified in the empirical section are funded. They are generated by a windfall that accrues to the state government. This is not an absolute windfall or the creation of resources. Rather this is a relative windfall that nets to zero (given the fixed effect normalization), so that a gain by one state must be accompanied by a loss in other states. Therefore the 'windfall experiment' is actually a form of transfer across states. In the rest of this section, I will explicitly model the 'windfall' as a transfer across states. Models (1) and (2) assume that the economy receiving the transfer is closed in all other respects, and hence omit any discussion of the transferring states. Models (3), (4) and (5) remove that assumption and include the effect of the transfer on both economies.

The transfer is received by the state government. In the models, the government can expend this transfer is one of two ways: wasteful government consumption or direct transfer payments to citizens. Having established these preliminaries, I can now consider the five models.[47]

[47] To make this discussion manageable, I consider only a subset of all possible models in this text. Here I consider static models. This omits the possibility that a government can save resources. Additionally, in a dynamic model where nominal interest rates are unresponsive to cross-state transfers, short-run nominal stickiness could impact income by reducing the real interest rate as in Christiano, Eichenbaum and Rebelo (2009). I also do not allow for complementarities between labor, public and private consumption (see Hall 2009 for a discussion). Similarly, I do not consider alternate forms of preferences which modify the standard wealth effect on labor supply, such as those in Greenwood, Hercowitz and Huffman (1988) and Jaimovich and Rebelo (2007). These extensions are considered in the online appendix.

Model 1: In an endowment economy, under the measurement definition from the empirical section, the income effect of windfall spending will be equal to the fraction of the transfer remitted to individuals.

Model 1 illustrates the accounting behind state personal income, the dependent variable in the empirical section. This fact highlights the important theoretical distinction between government consumption (generally modeled as dissipative) and transfer to households. Transfers alone have a direct effect on income in an endowment economy.[48]

The empirical exercises did not distinguish between these types of spending. Though some data on this division does exist[49], the distinctions drawn in the data do not necessarily match the relevant theoretical distinctions. For example, transferring income by increasing the salaries of public employees would register as a transfer payment in the model but not in the data. The empirical work strove to remain agnostic about the correct definition.

The model, however, cannot remain agnostic on this distinction. As noted above, only the remitted component of spending has a direct effect on income. The main finding in this section is that the standard model does not produce a multiplier greater than one, a fact that holds true for both types of spending. For expositional purposes, then, I proceed by considering only the effect of transfer spending, which generates the larger baseline income response. Comparable results for government consumption in this model are derived in the online appendix.

Model 2: In a closed economy with endogenous labor, income will increase by less than the fraction remitted to individuals. The multiplier will be less than one and there will be a reduction in hours worked.

This model confirms the preview at the beginning of this section. In order for the intratemporal optimality condition to hold after the transfer, consumption and labor must move in opposite directions. The wealth effect reduces the agent's labor supply and without any countervailing substitution effect, reduces output. The fall in output generates a multiplier below one, as in Mulligan (2010). Appendix B contains a formal proof.

[48]This is a statement about the accounting definition only in the endowment economy. When labor is endogenous, a government could wastefully spend its resources by hiring labor. This wasteful spending would be counted under personal income, though in the frictionless closed-economy model this crowds out an equal amount of private labor, producing no income effect.

[49]When using personal income excluding transfers (as measured by the BEA) in the baseline IV regression, the coefficient on spending is 2.30** (SE=1.0). The slightly higher number reflects a small decline in income from transfers, which include income maintenance transfers, unemployment benefits and Medicaid payments.

Model 3: In an open economy, the 'windfall experiment' raises income by less than one and reduces hours worked.

To make this statement, I first need to construct an open economy model with a transfer between states. The same basic setup will be used for Models (4) and (5) as well. For simplicity, I assume two symmetric states each of which produce a single state-specific traded good. I further assume that the entire transfer is remitted to the citizens, so that the states can be modeled as representative agents. The states have a fixed nominal exchange rate (to match the situation of a U.S. state when I consider nominal frictions), and I allow for home bias in consumption. To ensure determinancy in Models (4) and (5), when prices and wages are preset, I use a decreasing return to scale production function.

The model can be described in five equations that hold in analogous versions for each state. The full model (described in Appendix B) augments these equations with market clearing conditions and the requirement that transfers sum to zero. The equations are as follows.

Utility is given by:

$$U = \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{\eta}{1+\phi} H_t^{1+\phi}$$

where σ measures the curvature of utility from consumption, and ϕ^{-1} is the Frisch elasticity of labor supply.

Each state's consumption bundle is a CES aggregator of home and foreign goods:

$$C_t = \left(\gamma^{\frac{1}{\theta}} C_{H,t}^{1-\frac{1}{\theta}} + (1-\gamma)^{\frac{1}{\theta}} C_{F,t}^{1-\frac{1}{\theta}} \right)^{\frac{1}{1-\frac{1}{\theta}}}$$

where $\gamma \in (.5, 1)$ measures the preference for domestic goods and θ is the elasticity of substitution between goods. The state's budget constraint is given by:

$$P_t C_t = W_t H_t + Profits + T_t$$

where T_t is the transfer between states. I assume firms are owned domestically, so that profits and labor income sum to the value of domestic output.

Lastly, the model contains two price-setting equations:

$$P_{Ht} = \Omega_p \frac{W_t H_t}{H_t^{\alpha}}$$

$$W_t = \Omega_w P_t C_t^{\sigma} H_t(i)^{\phi}$$

I assume that prices are set at a constant markup over marginal costs and wages are set as a constant markup over the marginal rate of substitution. These assumptions can be easily micro-founded with a standard monopolistically competitive framework. In Models (4) and (5), these assumptions will be used to insert frictions in to the intratemporal optimality condition. In Model (3), I assume both markups (and profits) are set equal to zero.

I solve the model using log-linearization. Expressing the variables of interest in logs, I show in Appendix B that under mild conditions:

$$c_t = \frac{1}{2\left(1-\gamma\right)\left[1-\lambda\left(1-2\theta\gamma\right)\right]}\frac{dT}{\bar{PC}} > 0$$

$$h_t = \frac{2\gamma - 1 - \theta 4\gamma\left(1-\gamma\right)\lambda}{2\alpha\left(1-\gamma\right)\left[1-\lambda\left(1-2\theta\gamma\right)\right]}\frac{dT}{\bar{PC}} < 0$$

where $\lambda = \frac{\sigma\alpha+(2\gamma-1)(\phi+(1-\alpha))}{\left[\alpha+\theta(\phi+1-\alpha)-\theta(2\gamma-1)^2(\phi+1-\alpha)-\alpha(2\gamma-1)\right]}$, and $\frac{dT}{\bar{PC}}$ is the size of the transfer scaled to steady state output.

This result shows that an open-economy generalization of Model (2) does not change the basic response governed by the intratemporal condition.[50] The wealth effect of the transfer once again reduces labor supply and causes a fall in employment and a multiplier less than one. As intuition would suggest, the reduction in labor increases in σ and falls with ϕ.

Model 4: When prices are sticky, the 'windfall experiment' increases hours worked, producing a multiplier greater than one.

To model sticky prices, I now set the markup on prices $\Omega_p > 0$. When markups are non-zero, the intratemporal optimality condition becomes:

$$\left(\alpha - 1\right)h_t = \sigma c_t + \phi h_t + \mu_t^p$$

In this equation consumption and labor can rise together, provided that the price markup falls at the same time. In Appendix B, I solve for the response of consumption and output when prices are set prior to the transfer. The resulting expressions are:

$$c_t = \frac{1}{2\left(1-\gamma\right)}\frac{dT}{\bar{PC}} > 0$$

$$h_t = \frac{2\gamma-1}{2\alpha\left(1-\gamma\right)}\frac{dT}{\bar{PC}} > 0$$

Labor now rises following the transfer, indicating a multiplier greater than one. The mechanism behind this result is intuitive. In the presence of sticky prices, the transfer between states plays the role of shock to the money supply. The extra cash balances raise the demand for the domestic good. With sticky prices, producers meet this demand by increasing wages and reducing their markup. Falling markups induce a substitution effect that raises equilibrium output. Thus, in a model with pre-set prices and a wedge in

[50] This model also allows one to study the 'transfer problem' debate of Keynes (1929) and Ohlin (1929) discussed above. When consumers display home bias in consumption, the model predicts an appreciation in the terms of trade of the transfer recipient following the prediction of Keynes. This result is also derived in Appendix B.

the intratemporal condition, the multiplier is greater than one. This effect is stronger when the relative preference for the home good γ is larger.

Note that, in most models, sticky prices alone do not generate large multipliers (Hall 2009). This case differs from the standard one in that the spending shock is both financed and consists of monetary transfers to individuals. This means that there is no negative wealth effect on households, which would otherwise shift out their labor supply curve and reduce the fall in markups.

Model 5: When wages are sticky, the 'windfall experiment' again increased hours worked—producing a multiplier greater than one.

In this model, I once again set the price markup $\Omega_p = 0$ and instead let the wage markup $\Omega_w > 0$. Once again this inserts a wedge into the intratemporal optimality condition, allowing consumption and labor to move together.

In Appendix B, I solve for the response of consumption and output when wages are set before the transfer. The expressions for consumption and labor in Model (5) are:

$$c_t = \frac{[\alpha + \theta 4\gamma (1-\gamma)(1-\alpha)]}{[[\alpha + \theta 4\gamma (1-\gamma)(1-\alpha)] - (2\gamma - 1)[1-(2\gamma-1)(1-\alpha)]]} \frac{dT}{\bar{P}\bar{C}} > 0$$

and

$$h_t = \frac{(2\gamma - 1)}{[[\alpha + \theta 4\gamma (1-\gamma-\alpha+\alpha\gamma)] - (2\gamma - 1)[1-(2\gamma-1-\alpha 2\gamma+\alpha)]]} \frac{dT}{\bar{P}\bar{C}} > 0$$

As before, the addition of a nominal friction generates an increase in labor following the transfer. The mechanism now works differently. In the presence of home bias, the transfer raises demand for the home good. With sticky wages, labor is perfectly willing to accommodate the increased demand. This reduces the wage markup. The increase in labor income further raises personal income generating a multiplier greater than one.

Models (4) and (5) both predict a multiplier greater than one. Despite this, the two models generate different prediction about the effect of windfalls on real wages. At the moment it is impossible to test this effect empirically, due to the lack of state-level price data. I am currently in the process of constructing a state-level price series using data from the BLS & the ACCRA Cost of Living survey. I hope to explore this issue in future versions of the paper.

Again, in this exercise I've demonstrated the need to modify the standard intratemporal optimality condition in order to match the empirical results. I believe that this speaks to the general empirical relevance of that condition. Many papers, such as Chari, Kehoe and McGrattan (2004) have also shown that deviations from this condition ('the labor wedge') are needed to match national data. When this condition does not hold, it becomes far easier to generate large spending multipliers even in the tax-or-debt financed wasteful spending case (see Monacelli and Perotti 2008 for a discussion). To the extent that the empirical results suggest a meaningful distortions in the wage or product markets, this exercise speaks to the likelihood of finding large multipliers beyond the 'windfall experiment.'

9 "Is this Estimate of the Multiplier too High?"

The spending multiplier estimated in this paper is large relative to the conventional wisdom on the impact of state-level non-defense spending. However, a review of the literature reveals that it is difficult to support the conventional wisdom using the available empirical evidence.

The estimated income effect presented in this paper lies within the range of estimates presented in the VAR literature. In fact, Caldara and Kamps (2008) demonstrate that the three main identification techniques used in the VAR literature all deliver a multiplier of 2 at the three-to-four year horizon.

In contrast, the estimates in this paper are indeed higher than the estimates in the 'military events' literature. That literature, however, speaks directly only to wartime defense expenditures. There are a number of reasons to expect the mechanism to be different in those instances, including high capacity utilization, concurrent rationing and large fraction of expenditure being spent abroad. The papers in the military events literature that do look at non-defense expenditure (such as Barro and Redlick 2010) find large and comparable multipliers. Additionally, Hooker and Knetter (1997), who look at the effect of military spending on state-level outcomes, find large employment effects. Neither the VAR nor the military events literature provides reliable support to the claim that the multiplier on state-level peace-time non-defense spending is small.

Only a handful number of papers (including this one) have attempted to measure the income effect of state-level non-defense spending using instrumental variables. To my knowledge, the only other studies taking this approach are Clemens and Miran (2010), Chodorow-Reich et al. (2010) and Fishback and Kachanovskaya (2010). The identifying sources of variation differ considerably across these papers. Nevertheless, all three papers deliver impact multipliers very similar to the ones presented in this paper (baseline effects of 1.7 and 2.0 and 1.67 respectively). The results of all four existing IV studies are very similar, and all conclude that the multiplier on non-defense spending is large.

Indirect evidence on the size of the multiplier is given by estimates of the marginal propensity to consume, at least under the traditional Keynesian framework. Parker, Souleles, Johnson and McClelland (2009) estimate an MPC between 50-90% using data on the 2008 Economic Stimulus Payments. Johnson, Parker and Souleles (2006) estimate an MPC of roughly two-thirds over the six month horizon, using data on the 2001 income tax rebate. These estimates are too large to reconcile with the permanent income hypothesis.

They suggest credit-constrained or rule-of-thumb consumers, a feature that is consistent with models that generate large spending multipliers (Gali, Lópes-Salido, & Vallés 2005).

The evidence discussed above appears to support a large multiplier on non-defense spending generally. There are a number of reasons, though, to expect the multiplier studied in this paper to be larger than a multiplier on spending that must be financed by taxes or debt. Abstracting from these negative effects should increase the estimated effect on income. Additionally, as demonstrated in section 8, transfers can induce reinforcing movements in the terms-of-trade. Thus while the income and employment effects documented here may appear high at first blush, they actually are well within the range of estimates supported by the available data.

10 Conclusion

This paper estimates a fiscal multiplier using an instrument constructed from the windfall returns of state-administered pension plans. These returns are not explained by in-state bias or political considerations, indicating that the idiosyncratic returns earned by these plans are unrelated to in-state economic conditions. Weak market efficiency implies that these excess returns are not expected beforehand, and the large size of these plans ensures sizeable windfall shocks. Thus, the 'windfall' funds earned by state-run pension plans are exogenous, unexpected and large in magnitude.

In the first-stage results I show that these pension windfalls lead to large increases in state government spending. This relationship is robust across controls, specifications and constructions of the instrument. The effect is present in the cross-section and the time-series, and as evidenced by the placebo checks, is not due to a lack of specificity in the tests.

The second-stage regressions demonstrate that higher levels of state spending generate significant increases in income and employment. These effects are also robust across specifications, constructions of the instrument and the exclusion of individual states and years. Further, the effects are present across state borders, concentrated in non-traded industries, and larger during periods of economic 'slack'. The estimated multiplier of just over 2 might seem surprisingly high, but in fact is consistent with the empirical literature on the effect of non-defense state-level spending.

A theoretical exploration of this "windfall experiment" reveals that the standard macroeconomic model has difficulty matching these findings. However, the model was able to generate a multiplier greater than one when paired with frictions that drive a wedge between the marginal rate of substitution and the marginal product of labor. Other papers attempting to deliver large multipliers at the national level also require wedges in the intratemporal optimality condition. The empirical findings in this paper could be interpreted as evidence supporting these modeling assumptions.

This paper has only scratched the surface with regards to what can be accomplished with a strong instrument for fiscal policy. State pension windfalls can be used to explore the impact of government spending on everything from price-levels and real product wages to re-election rates. Additional information on the institutional details of these plans has the potential to multiply the number of instruments that can be constructed from these windfalls. These instruments could then be used to separately identify the effects of different types of spending. Work along of these dimensions has already begun.

These findings have important implications for policy. State level economic shocks are substantial. The large income and employment effects demonstrated here indicate that cross-state transfers could play a large role in smoothing these risks.

To close, I want to reiterate that the income and employment effects estimated here are not directly applicable to national tax-or-debt financed stimulus. Though I believe these results are indirectly informative about the general effectiveness of fiscal policy, there are important differences between that situation and the experiment in this paper. Caution should be used when extrapolating beyond the circumstances considered here.

References

Association, National Education (1998): *Characteristics of 100 Large Public Pension Plans.* Research Division.

Athanasoulis, S., and E. van Wincoop (2001): "Risksharing within the US: What Do Financial Markets and Fiscal Federalism Accomplish?," *Review of Economics and Statistics*, 83(4), 688–698.

Aubry, J.-P., A. H. Munnell, and D. Muldoon (2008): "The financial crisis and state/local defined benefit plans," Center for Retirement Research working paper.

Barrios, T., R. Diamond, G. W. Imbens, and M. Kolesar (2010): "Clustering, Spatial Correlations and Randomization Inference," NBER Working Paper No. w15760.

Barro, R., and C. Redlick (2010): "Macroeconomic Effects from Government Purchases and Taxes," mimeo.

Bartik, T. J. (1991): *Who Benefits from State and Local Economic Development Policies?* W. E. Upjohn Institute for Employment Research, Kalamazoo, MI.

Blanchard, O., and L. Katz (1992): "Regional Evolutions," *Brookings Papers on Economic Activity*, 23(1), 1–76.

Blanchard, O., and R. Perotti (2002): "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics*, 117(4), 1329–1368.

Bovbjerg, B. (2008): "State and local government pension plans: Current structure and funded states," United States Government Accountability Office, testimony before the Joint Economic Committee.

Brainard, K. (2008): *Public Fund Survey Summary of Findings for FY2007.* National Association of State Retirement Administrators.

Brown, J. R., J. Pollet, and S. J. Weisbenner (2009): "The Investment Behavior of State Pension Plans," mimeo.

Caldara, D., and C. Kamps (2008): "What are the effects of fiscal shocks? A VAR-based comparative analysis," European Central Bank Working Paper No. 877.

Chari, V., P. Kehoe, and E. McGrattan (2006): "Business Cycle Accounting," Federal Reserve Bank of Minneapolis Staff Report 328.

Chirinko, R. S., and D. J. Wilson (2009): "A state level database for the manufacturing sector: construction and sources," Working Paper Series, Federal Reserve Bank of San Francisco.

Chodorow-Reich, G., L. Feiveson, Z. Liscow, and W. G. Woolston (2010): "Does State Fiscal Relief During Recessions Work? Evidence from the American Recovery and Reinvestment Act," mimeo.

Christiano, Lawrence, Martin Eichenbaum, Sergio Rebelo (2009); "When is the Government Spending Multiplier Large?" mimeo.

Clemens, J., and S. Miran (2010): "The Effects of State Budget Cuts on Employment and Income," mimeo.

Cohen, L., J. Coval, and C. Malloy (2010): "Do Powerful Politicians Cause Corporate Downsizing?," mimeo.

Damodaran, A. (1991): *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset.* John Wiley & Sons, New York.

Fama, E. F., and J. D. Macbeth (1973): "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81(3), 607–636.

Fishback, P., and V. Kachanovskaya (2010): "In Search of the Multiplier for Federal Spending in the States During the New Deal," mimeo.

Gali, J., J. D. López-Salido, and J. Vallés "Understanding the Effects of Government Spending on Consumption, (Jordi Gali, J.D. López-Salido and J.Vallés), *Journal of the European Economics Association*, vol. 5, issue 1, 2007, 227–270.

Glaeser, E. L., and R. E. Saks (2004): "Corruption in America," Harvard Institute of Economic Research Discussion Paper No. 2043.

Gramlich, E. M. (1977): *Intergovernmental Grants: A Review of the Empirical Literature* pp. 219–239. Lexington Press, Lexington, MA.

Greenwood, Jeremy, Zvi Hercowitz and Gregory Huffman (1988): "Investment, Capacity Utilization and the Real Business Cycle," *American Economic Review* 78, 402-417, 1988.

Hall, Robert (2009): "By How Much Does GDP Rise if the Government Buys More Output?" *Brookings Papers on Economic Activity*, 2009, 2, pp. 183-231.

Hines, J. R., and R. H. Thaler (1995): "The Flypaper Effect," *Journal of Economic Perspectives*, 9(4), 217–226.

Hooker, M. A., and M. M. Knetter (1997): "The Effects of Military Spending on Economic Activity: Evidence from State Procurement Spending," *Journal of Money, Credit, and Banking*, 28(3), 400–421.

Jaimovich, Nir, and Sergio Rebelo. (2009): "Can News about the Future Drive the Business Cycle?" *American Economic Review*, 99(4): 10971118.

Johnson, D. S., J. A. Parker, and N. S. Souleles (2006): "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, 96(5), 1589–1610.

Keynes, J. M. (1949 [1929]): *The German Transfer Problem*, Blakiston Company, Philadelphia and Toronto.

Kraay, Aart (2010): "How Large is the Government Spending Multiplier? Evidence from World Bank Lending," mimeo.

Lane, P. R., and G. M. Milesi-Ferretti (2004): "The Transfer Problem Revisited: Net Foreign Assets and Real Exchange Rates," *The Review of Economics and Statistics*, 86(4), 841–857.

Monacelli, T., and R. Perotti (2008): "Fiscal Policy, Wealth Effects, and Markups," NBER Working Papers No. 14584.

Mountford, A., and H. Uhlig (2009): "What are the effects of fiscal policy shocks?," *Journal of Applied Econometrics*, 24(6), 960–992.

Mulligan, Casey (2010): "Simple Analytics and Empirics of the Government Spending Multiplier and Other Keynesian Paradoxes," mimeo.

Munnell, A. H., and A. Sundén (1999): "Investment Practices of State and Local Pension Funds: Implications for Social Security Reform," Center for Retirement Research at Boston College Working Paper, #2000-01.

Novy-Marx, R., and J. Rauh (2010): "The Crisis in Local Government Pensions in the United States," mimeo.

Ohlin, B. (1949 [1929]): *The Reparation Problem: A Discussion*, Blakiston Company, Philadelphia and Toronto.

Olea, P. M., and C. Pflueger (2010): "Is $F > 10$ enough? TSLS Weak Instrument Bias with Heteroskedasticity and Autocorrelation," mimeo.

Pappa, Evi, "The effects of fiscal shocks on employment and real wage," *International Economic Review*, Vol. 50, Issue 1, pp. 217–244, February 2009.

Parker, J. A., N. S. Souleles, D. S. Johnson, and R. McClelland (2009): "Consumer Spending and the Economic Stimulus Payments of 2008," mimeo.

Perotti, R. (2005): "Estimating the Effects of Fiscal Policy in OECD Countries," CEPR Discussion Papers.

Ramey, V. (2010): "Identifying Government Spending Shocks: It's All in the Timing," mimeo.

Rauh, Joshua (2006): "Investment and Financing Constraints: Evidence from the Funding of Corporate Pension Plans," *Journal of Finance* 61(1), 2006, 33–71.

Shoag, D. (2010): “The Investment Decisions of State-Administered Pension Plans,” mimeo.

State Budget Officers, The National Association of. (various): *The Fiscal Survey of States*. NASBO, Washington DC.

Thompson, S.B. (2009): “Simple Formulas for Standard Errors that Cluster by Both Firm and Time.” mimeo.

U.S. Census Bureau, P. D. (1990): *Journey To Work and Place Of Work*.

Zorn, P. (1991): *Survey Report: 1991 Survey of State and Local Government Employee Retirement Systems*. Government Finance Officers Association, Chicago, IL.

——— (1993): *Survey Report: 1993 Survey of State and Local Government Employee Retirement Systems*. Government Finance Officers Association, Chicago, IL.

——— (1995): *Survey Report: 1995 Survey of State and Local Government Employee Retirement Systems*. Government Finance Officers Association, Chicago, IL.

——— (1997): *Survey Report: 1997 Survey of State and Local Government Employee Retirement Systems*. Government Finance Officers Association, Chicago, IL.

Appendix A: Timing and Sizing the Fall in Contributions

The text demonstrates that the excess return earned by a state-administered pension plan is a strong predictor of state government spending. The hypothesized channel is that these returns alter the contributions states make to these pension systems, freeing or constraining the resources available for other types of spending. In this section, I explore this channel, and the considerable heterogeneity in the pension budgeting process. The main conclusion in this section is that the cumulative effect of return shocks on state contributions is comparable to the associated increase in spending. The fall in contributions (in the Census data) is phased in over a considerably longer horizon than the increase in spending. This suggests states capitalize on lower expected costs in the future. This is consistent with the fact, shown above, that states with weaker balanced budget restrictions have larger immediate spending responses.

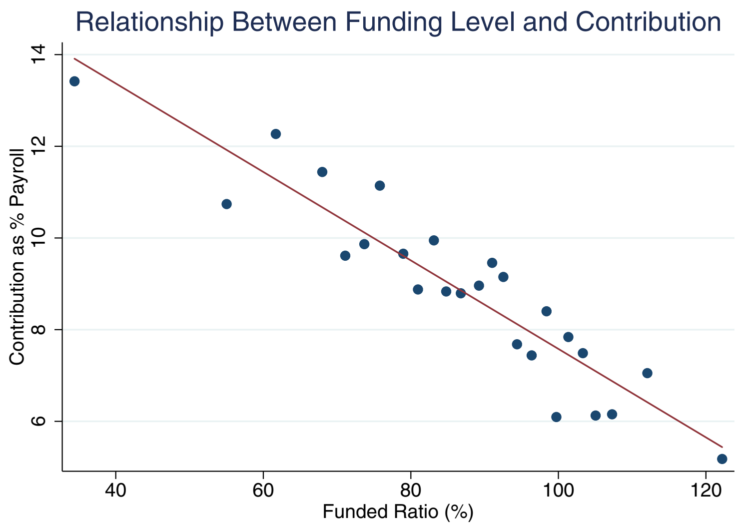
Using my new data on the market value of plan assets, I also show that the growth in asset holdings following high returns dies out more quickly than the measured reduction in contributions. The fading out in the plan asset data corresponds closely to the timing of the spending response demonstrated above. This incongruity with the Census data is puzzling, and might indicate that states find a way to tap pension fund windfalls in a more complicated fashion.

In any case, it's important to make two points before proceeding. The first is that, under all measures, the contributions made to state-administered systems are lower when the plan is better funded. This is evident in both my data and the Census data. I also confirmed this fact using data from a third independent source, the Wisconsin Legislative Council's Comparative Retirement Studies. The figure below plots the relationship between funding status and contributions in that data.

The second point is that whether or not there is a reduction in contributions in the short-run is irrelevant to the identification assumptions underlying the instrument. States manage many funds and may invest them similarly. A windfall earned by the pension fund may be accompanied by similar excess returns from many funds. While this point may cast some doubt on the 'flypaper' estimates in the first section, it would not affect the validity of the second-stage results. Similarly, naiveté on the part of state officials may generate a larger spending response than the fall in contributions. This too does not bias the paper's fundamental results.

Having made this point I now discuss some of the heterogeneity in plan budgeting processes. Within the general framework described in the text, there is a substantial amount of institutional variety.

Ninety-four percent of the plans in my sample for which data was available in the 2001 PENDAT survey conduct an annual actuarial valuation of their funding status, with the remaining six percent conducting a valuation every other year. Eighty-two percent of plans in this survey report that the annual contribution is



Note: These data are from the Wisconsin Legislative Council's Comparative Retirement Study. The graph plots the percentage of payroll contributed to the state administered plan against the funding status of the plan.

set directly by the valuation study. Actuarial smoothing techniques also differ across plans. Between seven and twelve percent of the plans matched to my data use market values in their actuarial valuations, with the remaining plans smoothing returns over three to five years. The amortization horizon similarly varies across plans. In the PFS sample approximately 17% of plans had horizons at or below twenty years, 69% of plans had horizons between twenty one and thirty years and 14% percent of plans had horizons exceeding thirty years.[51] These actuarial mechanisms are designed to smooth return shocks and reduce fiscal pressure on the state.

The 1998 PENDAT survey contains information on the frequency of contributions to retirement funds, as well as the source of funding.

Frequency of Contributions	**Fraction of Plans**
Yearly	8.97%
Quarterly/ Semi-Annually	15.38%
Monthly	33.33%
Pay	42.31%

Source of Funds	**Fraction of Plans**
General Fund	82.05%
State Appropriation	10.26
Special Tax	6.41%
Other	1.28%

[51] GASB regulations set a maximum amortization period of 30 years for all open plans. These regulations are non-binding, but generally followed by state administered systems.

The Census' Survey of Public Employees Retirement Systems has data on employer contributions to these plans. Under the hypothesized mechanism, a windfall shock should free up resources by reducing these required contributions.

To test this mechanism and explore the timing of the response, I regress per-capita plan-level contributions on a number of lags of the excess funds variable. The cumulative response is measured by summing the recovered coefficients. The regression uses a flexible lag length (chosen to minimize the AIC), and includes controls for time and plan. Additional details are provided below Table A1.

The regressions indicate that pension windfalls do reduce contributions. The initial drop in contributions per dollar of windfall is 6¢, in line with one would expect given the stated smoothing mechanism. The reduction in contributions grows over time, however, and reaches levels comparable to the increase in spending at the ten year horizon. This indicates that states are able to shift forward the eventual reduction in contributions and finance spending at shorter horizons. This is consistent with the fact that the spending response is larger for states with weak balanced budget amendments.

Table A1: The Effect of Windfalls on Employer Contributions

Cumulative Drop in Contributions	Contrib_t	$\text{Contrib}_{t,t+3}$	$\text{Contrib}_{t,t+5}$	$\text{Contrib}_{t,t+7}$	$\text{Contrib}_{t,t+10}$
Baseline:					
Excess Funds_{t-1}	-0.0658*	-0.127*	-0.234**	-0.426**	-0.517***
	(0.386)	(0.069)	(0.10)	(0.182)	(0.206)
Controls Benchmark Returns:					
Excess Funds_{t-1}	-0.060*	-0.110*	-0.184**	-0.356**	-0.425**
	(0.035)	(0.065)	(0.087)	(0.157)	(0.194)

Notes: The dependent variable is the level of real per capita contribution at the plan level. The table displays the of the coefficients on the lags of per capita excess returns. The standard errors are the Wald test of the linear combination, and are clustered by state. The regressions include year include a control for the number of active members in the plans. They also include unreported year dummies multiplied by the initial sample plan contribution to control for time effects. Plan fixed effects are again included but not reported. The lag length used in these specifications is 12, and was chosen tom minimizing the AIC.

Surprisingly, however, the delayed fall in contributions is not present in the plan-level asset data. In this data, a $1 windfall initially increases pension holdings by approximately $1. This increase in system assets associated with an extra dollar of returns dies out fairly rapidly, though, and at approximately the speed and inverse size of the increase in spending. This is demonstrated in Table A2 below, which uses a similar specification to the one used in Table A1. In this table, an extra dollar of returns seems to increase the long run value of system assets by about 45 cents. This closely matches the long run total spending increases of approximately 65 cents. The incongruity between these two data sources is puzzling, but as explained before, does not challenge the identification assumptions. Future research will explore the mechanism underlying this discrepancy.

Table A2: The Effect of Returns on Retirement System Assets

	(1)
	Plan Assets$_t$
Investment Return$_t$	1.08***
	(.11)
Investment Return$_{t-1}$	0.99***
	(.07)
Investment Return$_{t-2}$	0.72***
	(.08)
Investment Return$_{t-3}$	0.54***
	(.08)
Investment Return$_{t-4}$	0.44***
	(.05)
Investment Return$_{t-5}$	0.47***
	(.14)

Notes: Standard Errors are clustered by system. Regressions weighted by Portfolio Size. Regressions include unreported year dummies multiplied by the initial sample plan contribution to control for time effects. Plan fixed effects are also not reported.

Appendix B : Full Versions of Model (3) and (4), and Proofs

Model 2:

Proof: Let U= U(C,N) and Uc>0, Ucc<0 and Un<0 Unn>0, Fn>0 , Fnn<0

Then the FOC, when maximizing subject to the budget constraint is

$$U_c\left(F\left(N\right)+\theta,\,N\right)F_n=U_n(F\left(N\right)+\theta,\,N)$$

Taking the total derivative with respect to theta, gives

$$\left(U_{cc}\frac{dN}{d\theta}+U_{cc}\right)F_n+U_cF_{nn}\frac{dN}{d\theta}=U_{nn}\frac{dN}{d\theta}$$

So collecting terms gives:

$$\frac{dN}{d\theta}=\frac{U_{cc}F_n}{U_{nn}-U_cF_{nn}-U_{cc}F_n}\leq 0$$

and

$$\frac{d[F\left(N\right)+\theta]}{d\theta}=\frac{U_{cc}F_n^2}{U_{nn}-U_cF_{nn}-U_{cc}F_n}+1\geq 0$$

Model 3:

Review of the Model

Firm Block

Competitive Firms:

$$P_{Ht}H_t^{\alpha}=W_tH_t$$

(the markup $\Omega_p=0$)

Household Block

$$U=\frac{C_t^{1-\sigma}}{1-\sigma}-\frac{\eta}{1+\phi}H_t^{1+\phi}$$

$$C_t=\left(\gamma^{\frac{1}{\theta}}C_{H,t}^{1-\frac{1}{\theta}}+(1-\gamma)^{\frac{1}{\theta}}C_{F,t}^{1-\frac{1}{\theta}}\right)^{\frac{1}{1-\frac{1}{\theta}}}$$

$$P_t=\left(\gamma P_{H,t}^{1-\theta}+(1-\gamma)P_{F,t}^{1-\theta}\right)^{\frac{1}{1-\theta}}$$

$$P_tC_t=W_tH_t+T_t$$

Household First Order Conditions

$$C_{Ht}=\gamma\left(\frac{P_{Ht}}{P_t}\right)^{-\theta}C_t$$

$$C_{Ft}=\gamma\left(\frac{P_{Ht}}{P_t}\right)^{-\theta}C_t$$

Wage-Setting

$$W_t=P_tC_t^{\sigma}H_t^{\phi}$$

(the markup $\Omega_w=0$)

MARKET CLEARING

$$H_t^{\alpha}=\gamma\left(\frac{P_{Ht}}{P_t}\right)^{-\theta}C_t+\gamma\left(\frac{P_{Ht}}{P_t^*}\right)^{-\theta}C_t^*$$

The Log-Linear Equations are:

Labor Supply

$$w_t-p_t=\sigma c_t+\phi h_t \tag{1}$$

$$w_t^*-p_t^*=\sigma c_t^*+\phi h_t^* \tag{2}$$

Market Clearing

$$\alpha h_t=-\theta p_{Ht}+\gamma c_t+(1-\gamma)\,c_t^*+\gamma\theta p_t+(1-\gamma)\,\theta p_t^* \tag{3}$$

$$\alpha h_t^*=-\theta p_{F,t}^*+(1-\gamma)c_t+\gamma c_t^*+(1-\gamma)\,\theta p_t+\gamma\theta p_t^* \tag{4}$$

Price Indices

$$p_t=\gamma p_{Ht}+(1-\gamma)\,p_{F,t}^* \tag{5}$$

$$p_t^*=(1-\gamma)p_{Ht}+\gamma p_{F,t}^* \tag{6}$$

Profit Maximizing

$$p_{Ht}=w_t+(1-\alpha)\,h_t \tag{7}$$

$$p_{Ft}^*=w_t^*+(1-\alpha)h_t^* \tag{8}$$

$$c_t=(p_{Ht}-p_t)+\alpha h_t+\frac{dT}{\bar{P}C} \tag{9}$$

$$c_t^*=\left(p_{F,t}^*-p_t^*\right)+\alpha h_t^*+-\frac{dT}{\bar{P}C} \tag{10}$$

(1)-(2) generates

$$h_t-h_t^*=\frac{(w_t-w_t^*)-(p_t-p_t^*)-\sigma\,(c_t-c_t^*)}{\phi}$$

Subbing in (7) and (8), I get

$$h_t-h_t^*=\frac{\left(p_{Ht}-p_{F,t}^*\right)-(1-\alpha)\,(h_t-h_t^*)-(p_t-p_t^*)-\sigma\,(c_t-c_t^*)}{\phi}$$

$$\left[\frac{\phi+(1-\alpha)}{\phi}\right]h_t-h_t^*=\frac{\left(p_{Ht}-p_{F,t}^*\right)-(p_t-p_t^*)-\sigma\,(c_t-c_t^*)}{\phi}$$

$$h_t - h_t^* = \frac{\left(p_{Ht} - p_{F,t}^*\right) - (p_t - p_t^*) - \sigma\,(c_t - c_t^*)}{\phi + (1-\alpha)}$$

(3)-(4) generates

$$h_t - h_t^* = \frac{1}{\alpha}\left[-\theta\left(p_{Ht} - p_{F,t}^*\right) + (2\gamma - 1)(c_t - c_t^*) + \theta\,(2\gamma-1)(p_t - p_t^*)\right]$$

Combining terms

$$[\phi + (1-\alpha)]\left[-\theta\left(p_{Ht} - p_{F,t}^*\right) + (2\gamma-1)(c_t - c_t^*) + \theta\,(2\gamma-1)(p_t - p_t^*)\right] = \\ \alpha\left[\left(p_{Ht} - p_{F,t}^*\right) - (p_t - p_t^*) - \sigma\,(c_t - c_t^*)\right]$$

Using (5) and (6) gives

$$p_t - p_t^* = (2\gamma-1)(p_{Ht} - p_{F,t}^*)$$

So I get

$$\left[\alpha + \theta\,(\phi + 1 - \alpha) - \theta\,(2\gamma-1)^2(\phi + 1 - \alpha) - \alpha(2\gamma-1)\right]\left(p_{Ht} - p_{F,t}^*\right) = \\ [\sigma\alpha + (2\gamma-1)(\phi + (1-\alpha))](c_t - c_t^*)$$

$$\left(p_{Ht} - p_{F,t}^*\right) = \frac{[\sigma\alpha + (2\gamma-1)(\phi + (1-\alpha))]}{\left[\alpha + \theta\,(\phi + 1 - \alpha) - \theta\,(2\gamma-1)^2(\phi + 1 - \alpha) - \alpha(2\gamma-1)\right]}(c_t - c_t^*)$$

$$\left(p_{Ht} - p_{F,t}^*\right) = Terms\,of\,Trade = \lambda\,(c_t - c_t^*)$$

where $\lambda = \dfrac{[\sigma\alpha + (2\gamma-1)(\phi + (1-\alpha))]}{\left[\alpha + \theta\,(\phi + 1 - \alpha) - \theta\,(2\gamma-1)^2(\phi + 1 - \alpha) - \alpha(2\gamma-1)\right]}$

Using the above expression for the terms of trade, I get:

$$p_t - p_t^* = Real\,Exchange\,Rate = (2\gamma-1)\lambda\,(c_t - c_t^*)$$

I can derive an expression for the movement in output:

$$h_t - h_t^* = \frac{1}{\alpha}\left[-\theta\lambda\,(c_t - c_t^*) + (2\gamma-1)(c_t - c_t^*) + \theta\,(2\gamma-1)^2\lambda\,(c_t - c_t^*)\right]$$

Now I subtract (9) from (10)

$$\begin{aligned}
(c_t - c_t^*) &= \left(p_{Ht} - p_{F,t}^*\right) - (p_t - p_t^*) + \alpha\,(h_t - h_t^*) + \frac{2dT}{\bar{P}C} \\
(c_t - c_t^*) &= \lambda\,(c_t - c_t^*) - (2\gamma-1)\lambda\,(c_t - c_t^*) + -\theta\lambda\,(c_t - c_t^*) \\
&\quad + (2\gamma-1)(c_t - c_t^*) + \theta\,(2\gamma-1)^2\lambda\,(c_t - c_t^*) + \frac{2dT}{\bar{P}C} \\
2\,(1-\gamma)(c_t - c_t^*) &= 2\,(1-\gamma)\,\lambda\,(c_t - c_t^*) - 4\theta\gamma\,(1-\gamma)\,\lambda\,(c_t - c_t^*) + \frac{2dT}{\bar{P}C} \\
2\,(1-\gamma)\,[1 - \lambda\,(1 - 2\theta\gamma)]\,(c_t - c_t^*) &= \frac{2dT}{\bar{P}C} \\
(c_t - c_t^*) &= \frac{1}{2\,(1-\gamma)\,[1 - \lambda\,(1 - 2\theta\gamma)]}\frac{2dT}{\bar{P}C}
\end{aligned}$$

This is greater than zero if $\theta > \frac{1}{2\gamma}$

Using this result, I can determine what happens to labor.

$$h_t - h_t^* = \frac{1}{\alpha}\left[(2\gamma - 1) - \theta 4\gamma (1-\gamma)\lambda\right](c_t - c_t^*)$$

Plugging in I get:

$$h_t - h_t^* = \frac{1\left[(2\gamma - 1) - \theta 4\gamma (1-\gamma)\lambda\right]}{\alpha\left(2(1-\gamma)\left[1-\lambda(1-2\theta\gamma)\right]\right)} \frac{2dT}{\bar{P}\bar{C}}$$

When $\theta > \frac{1}{2\gamma}$, then the denominator is positive. So I consider the numerator. I get

$$(2\gamma - 1) - \frac{2(1-\gamma)\left[\sigma\alpha + (2\gamma-1)(\phi + (1-\alpha))\right]}{\left[2\alpha(1-\gamma) + \theta\phi + \theta - \theta\alpha - \theta\left(4\gamma^2 - 4\gamma - 1\right)(\phi + 1 - \alpha)\right]}$$

$$(2\gamma - 1) - \frac{2(1-\gamma)\left[\sigma\alpha + (2\gamma-1)(\phi + 1 - \alpha)\right]}{\left[2\alpha(1-\gamma) + 4\theta\gamma(1-\gamma)(\phi + 1 - \alpha)\right]}$$

$$(2\gamma - 1) - \frac{\left[\sigma\alpha + (2\gamma-1)(\phi + 1 - \alpha)\right]}{\left[\alpha + 2\theta\gamma(\phi + 1 - \alpha)\right]}$$

Which is negative for $\sigma > (2\gamma - 1)$. Labor falls under these conditions. These conditions are assumed in much of the literature.

Finally, I consider the effect on the terms of trade, in reference to the transfer problem debate between Keynes and Ohlin. Combining the expression for the terms of trade and the movement in relative consumption, we get:

$$\left(p_{Ht} - p_{F,t}^*\right) = Terms\,of\,Trade = \frac{\lambda}{2(1-\gamma)\left[1-\lambda(1-2\theta\gamma)\right]} \frac{2dT}{\bar{P}\bar{C}}$$

which under our conditions assumes the sign of λ, which itself is positive under the conditions assumed.

Model 4

In the sticky price case, a number of the equations used in Model 3 no longer apply. The remaining equations in the model are:

Labor Supply

$$w_t = \sigma c_t + \phi h_t \tag{11}$$

$$w_t^* = \sigma c_t^* + \phi h_t^* \tag{12}$$

Market Clearing

$$\alpha h_t = \gamma c_t + (1-\gamma) c_t^* \tag{13}$$

$$\alpha h_t^* = (1-\gamma) c_t + \gamma c_t^* \tag{14}$$

$$c_t = \alpha h_t + \frac{dT}{\bar{P}\bar{C}} \tag{15}$$

$$c_t^* = \alpha h_t^* + -\frac{dT}{\bar{P}\bar{C}} \tag{16}$$

Equations (5) and (6) assume full domestic ownership of firms, so that output is equal to wage income plus profits. I allow profits to be negative.

From these equations one can see that the wage setting equation does not determine the real variables in the economy. Subtracting (4) from (3) one derives

$$\alpha(h_t^* - h_t) = (2\gamma - 1)(c_t - c_t^*)$$

Subtracting (6) from (5), combining equations, and using the symmetry across states produces the desired result.

Model 5:

Now I consider the sticky wage case. Labor Supply is demand-determined so equations (1) and (2) from Model (3) no longer apply. The remaining equations are:

Market Clearing

$$\alpha h_t = -\theta p_{Ht} + \gamma c_t + (1-\gamma)\, c_t^* + \gamma\theta p_t + (1-\gamma)\,\theta p_t^* \tag{17}$$

$$\alpha h_t^* = -\theta p_{F,t}^* + (1-\gamma) c_t + \gamma c_t^* + (1-\gamma)\,\theta p_t + \gamma\theta p_t^* \tag{18}$$

Price Indices

$$p_t = \gamma p_{Ht} + (1-\gamma)\, p_{F,t}^* \tag{19}$$

$$p_t^* = (1-\gamma) p_{Ht} + \gamma p_{F,t}^* \tag{20}$$

Profit Maximizing

$$p_{Ht} = \bar{w} + (1-\alpha)\, h_t \tag{21}$$

$$p_{Ft}^* = \bar{w} + (1-\alpha) h_t^* \tag{22}$$

$$c_t = (p_{Ht} - p_t) + \alpha h_t + \frac{dT}{\bar{P}\bar{C}} \tag{23}$$

$$c_t^* = \left(p_{F,t}^* - p_t^*\right) + \alpha h_t^* + -\frac{dT}{\bar{P}\bar{C}} \tag{24}$$

Combining (1) and (2), I get

$$h_t - h_t^* = \frac{1}{\alpha}\left[-\theta\left(p_{Ht} - p_{F,t}^*\right) + (2\gamma-1)(c_t - c_t^*) + \theta\,(2\gamma-1)^2\left(p_{Ht} - p_{F,t}^*\right)\right]$$

$$h_t - h_t^* = \frac{1}{\alpha}\left[(2\gamma-1)(c_t - c_t^*) - \theta 4\gamma\,(1-\gamma)\left(p_{Ht} - p_{F,t}^*\right)\right]$$

Now from the pricing conditions :

$$p_{Ht} - p^*_{Ft} = (1-\alpha)\left(h_t - h^*_t\ \right)$$

$$\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right]\left(h_t - h^*_t\right) = \left(2\gamma - 1\right)\left(c_t - c^*_t\right)$$

$$\left(h_t - h^*_t\right) = \frac{\left(2\gamma-1\right)}{\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right]}\left(c_t - c^*_t\right)$$

$$p_t - p^*_t = (2\gamma-1)(p_{Ht} - p^*_{F,t})$$

Plugging these in to the Budget Constraints, I find

$$c_t - c^*_t = \left(p_{Ht} - p^*_{F,t}\right) - \left(p_t - p^*_t\right) + \alpha(h_t - h^*_t) + \frac{2dT}{\bar{P}\bar{C}}$$

$c_t - c^*_t =$

$$\left[\frac{\left(2\gamma-1\right)\left(1-\alpha\right)}{\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right]} + \frac{\alpha\left(2\gamma-1\right)}{\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right]} - \frac{\left(2\gamma-1\right)^2\left(1-\alpha\right)}{\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right]}\right]\left(c_t - c^*_t\right) + \frac{2dT}{\bar{P}\bar{C}}$$

$$\left[1 - \frac{\left(2\gamma-1\right)\left[1-\left(2\gamma-1\right)\left(1-\alpha\right)\right]}{\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right]}\right]\left(c_t - c^*_t\right) = \frac{2dT}{\bar{P}\bar{C}}$$

$$c_t - c^*_t = \frac{\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right]}{\left[\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right] - \left(2\gamma-1\right)\left[1-\left(2\gamma-1\right)\left(1-\alpha\right)\right]\right]}\frac{2dT}{\bar{P}\bar{C}}$$

Plugging this in, I get:

$$\left(h_t - h^*_t\right) = \frac{\left(2\gamma-1\right)}{\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right]}\frac{\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right]}{\left[\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right] - \left(2\gamma-1\right)\left[1-\left(2\gamma-1\right)\left(1-\alpha\right)\right]\right]}\frac{2dT}{\bar{P}\bar{C}}$$

$$\left(h_t - h^*_t\right) = \frac{\left(2\gamma-1\right)}{\left[\left[\alpha + \theta 4\gamma\left(1-\gamma-\alpha+\alpha\gamma\right)\right] - \left(2\gamma-1\right)\left[1-\left(2\gamma-1-\alpha 2\gamma+\alpha\right)\right]\right]}\frac{2dT}{\bar{P}\bar{C}}$$

Where again $\lambda = \dfrac{\left[\sigma\alpha + \left(2\gamma-1\right)\left(\phi + \left(1-\alpha\right)\right)\right]}{\left[2\alpha\left(1-\gamma\right) + \theta\left(\phi-\alpha\right) + 4\theta\gamma\left(1-\gamma\right)\right]}$

The numerator in both cases is clearly positive. So I need only check the denominator, which is the same in both cases. This means that the sign of the two responses will be the same.

I reduce the expression for consumption:

$$\frac{\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right]}{\left[\left[\alpha + \theta 4\gamma\left(1-\gamma\right)\left(1-\alpha\right)\right] - \left(2\gamma-1\right)\left[1-\left(2\gamma-1\right)\left(1-\alpha\right)\right]\right]}$$

Assuming the other conditions bind:

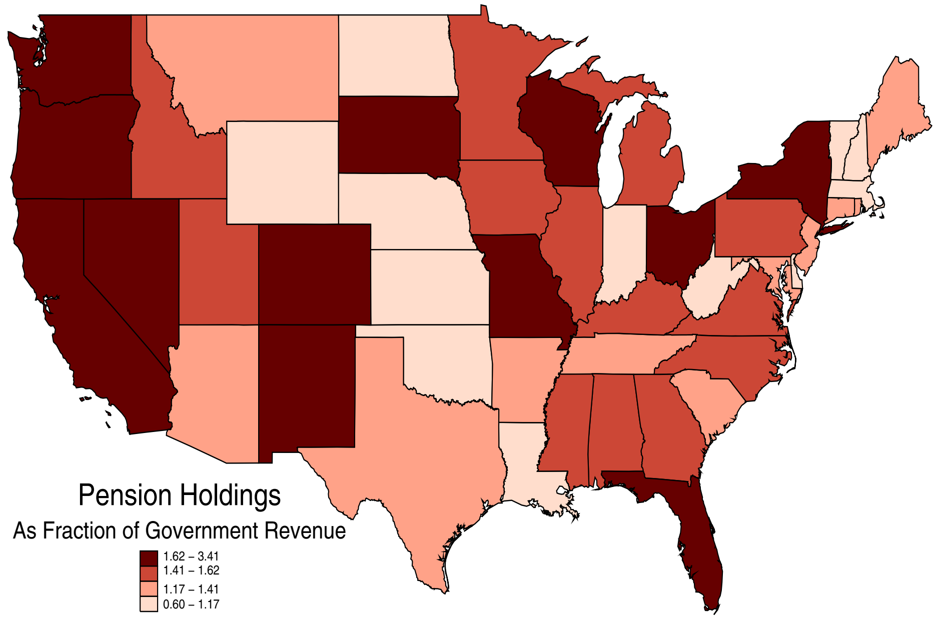
$$\left[\alpha + 2\left(1-\gamma\right)\left(1-\alpha\right)\right] - \left(2\gamma-1\right) + \left(2\gamma-1\right)^2\left(1-\alpha\right)$$

This expression is positive when alpha is greater than zero and the previous two conditions are met.

$$\left[\alpha + 2\left(1-\gamma\right)\left(1-\alpha\right)\right] - \left(2\gamma-1\right) + \left(2\gamma-1\right)^2\left(1-\alpha\right)$$

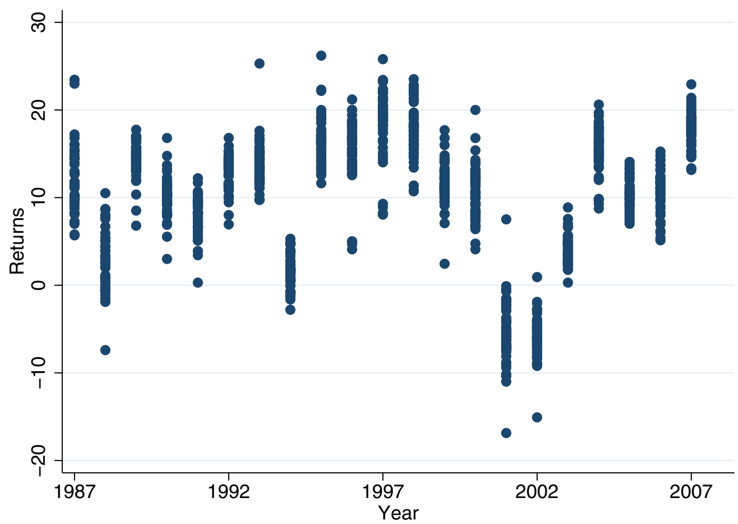
So both consumption and output rise following the transfer when wages are preset.

Figure 1: State-Administered Pension Plan Assets Relative to State Government Revenue.



Note: This figure graphs the ratio of state-administered pension assets to state government revenue in 2008. Asset holdings range from 60% to 341% of revenue. The figure demonstrates the variation across states in the relative size of state pension plans. The relative size of state pension funds does not appear to be distributed along geographic or demographic lines. Data are from the Census Bureau's State Government Finances and the State and Local Government Employees Retirement Systems surveys.

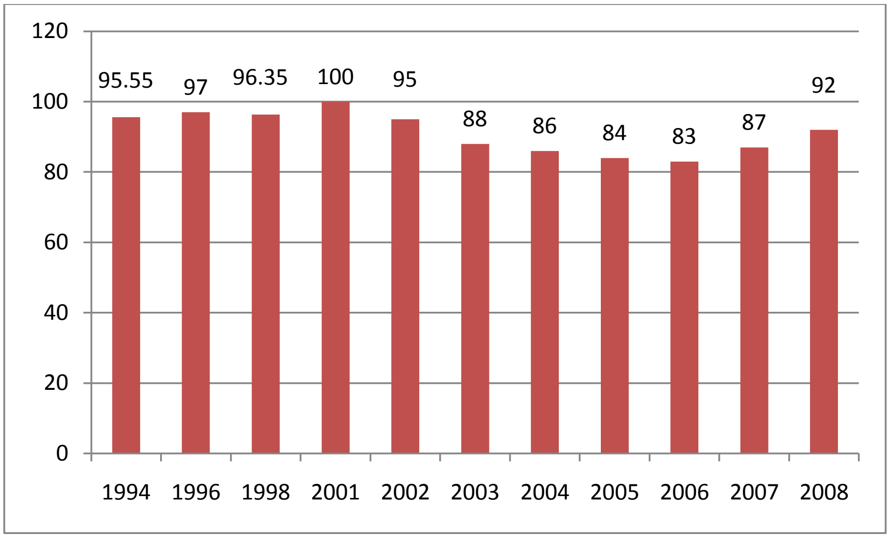
Figure 2: Distribution of Returns



Year	1988	1993	1998	2003	2008
Standard Deviation in Percentage Points	3.52	2.48	2.84	1.60	1.96

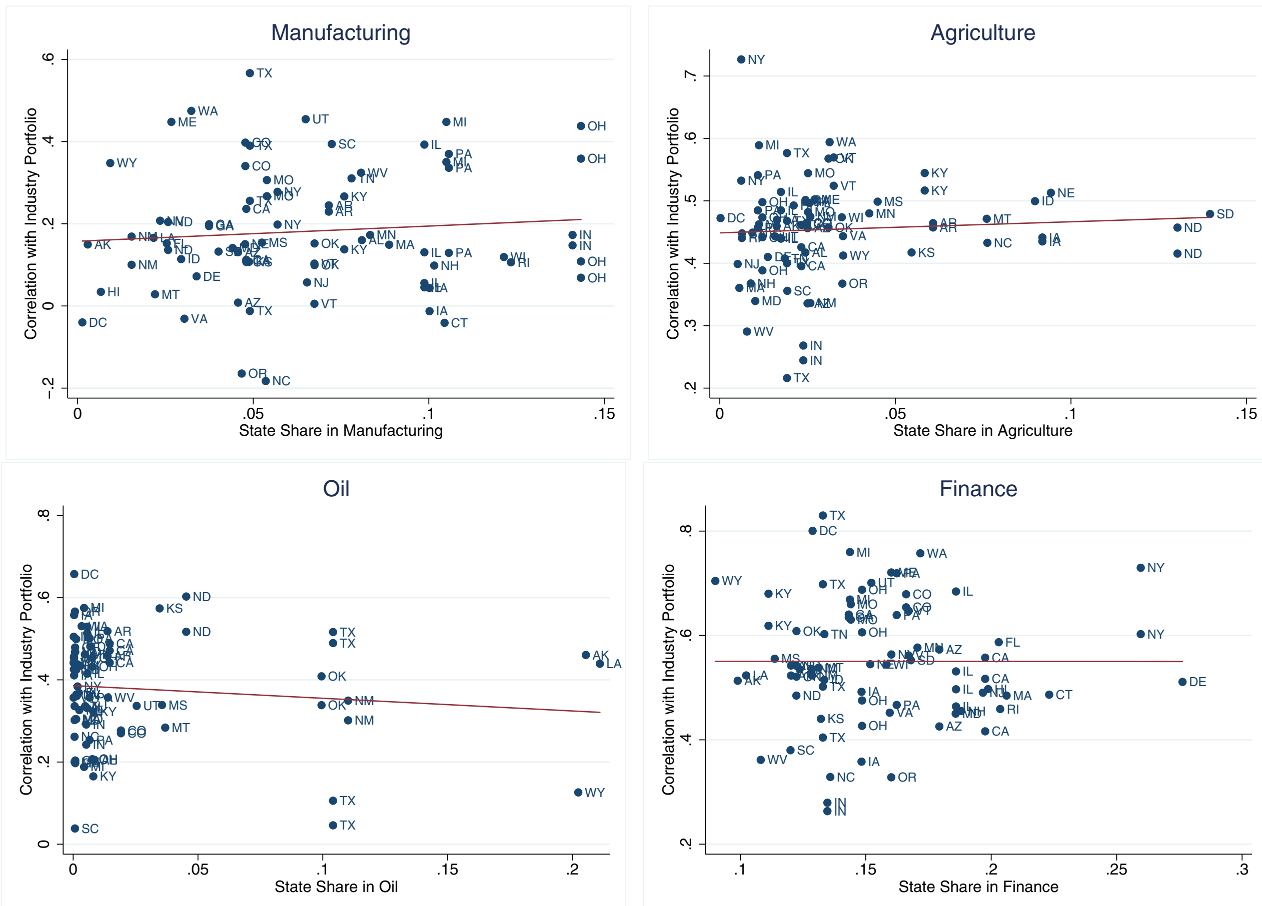
Note: The figure plots annual returns for plans with fiscal years ending in June. Both the vertical axis and the table are measured in hundreds of basis points. The table records the within-year standard deviation in returns across plans. The figure and the table demonstrate the considerable variation in returns across plans and time. See the text and the online appendix for a description of the data.

Figure 3: Percent of the ARC Contributed Over Time



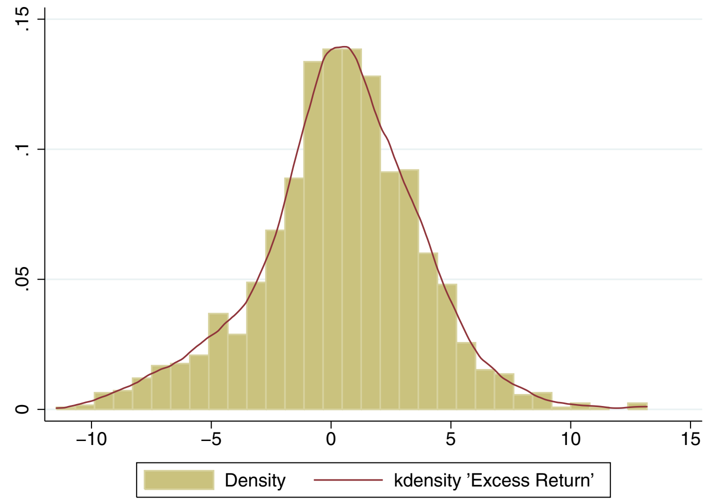
Note: The Actuarially-Required Contribution (ARC) is the contribution determined necessary to fund a plan's ongoing incurred liability and amortize its unfunded liability. The ARC is a function of the system's investment returns. This figure plots the average percentage of the ARC contributed by state administered plans. It shows that the ARC has considerable influence state contributions. Data from 2001-2008 are from Munnell, Aubry and Quinby (2010). Data from 1994-1998 are calculated by the author using data from the PENDAT surveys (Zorn, 1995, 1997, 1999). For these calculations, plans contributing more than the ARC were recorded as having contributed 100%.

Figure 4: Correlation of State Plans and Industry vs. % of State Industrial Share



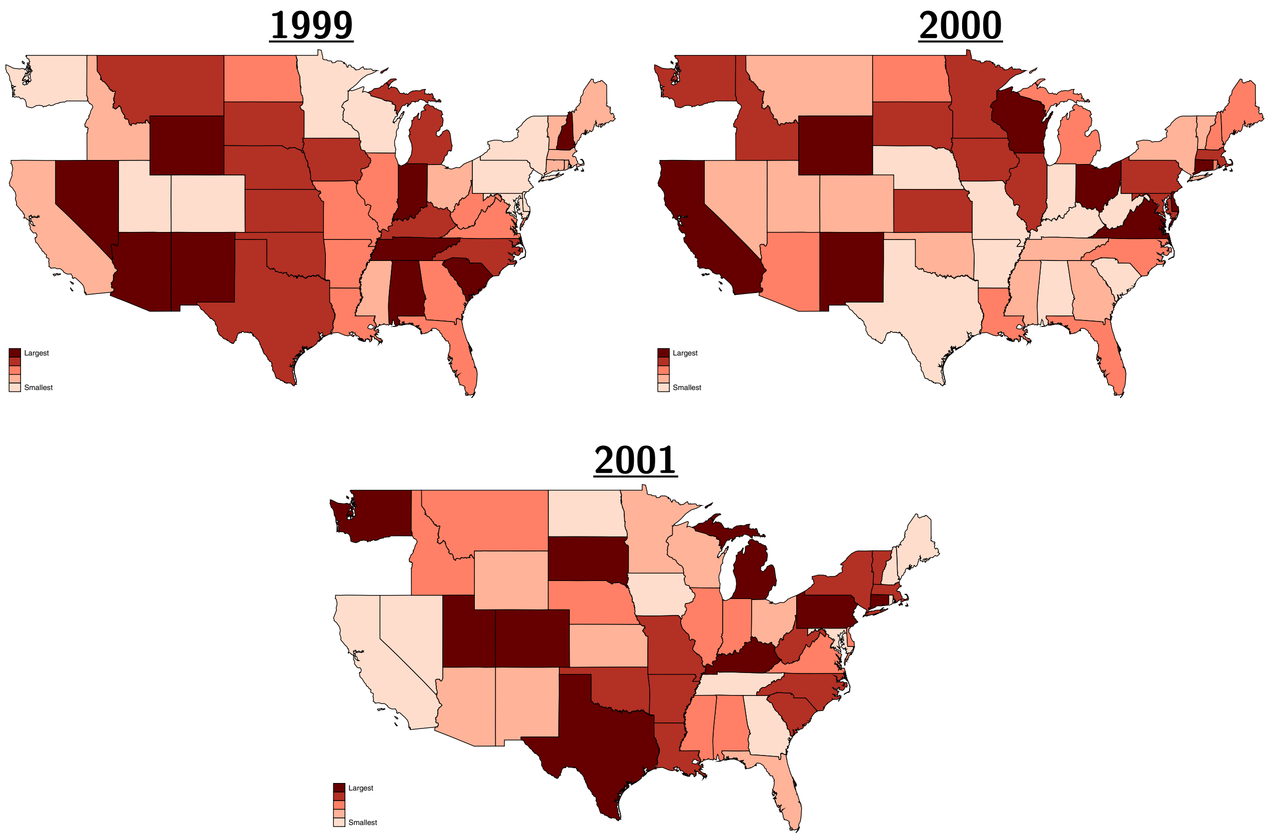
Note: This figure explores the correlation between state pension returns with state industries. The vertical axis measures the correlation between the plan's investment returns and the Fama-French industry return. The horizontal axis measures the share of state GSP produced in that industry. Data on state GSP by industry was taken from the BEA, and averaged over the years 1987-2008. Data on industry returns are from Kenneth French's data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). This figure shows that state plans do not have a higher correlation with industry portfolios when state economic activity is concentrated in that industry. This is evidence against the hypothesis that states over-invest in state or in state-specific industries.

Figure 5: Distribution of Plan Return minus the Benchmark Return: Baseline Instrument



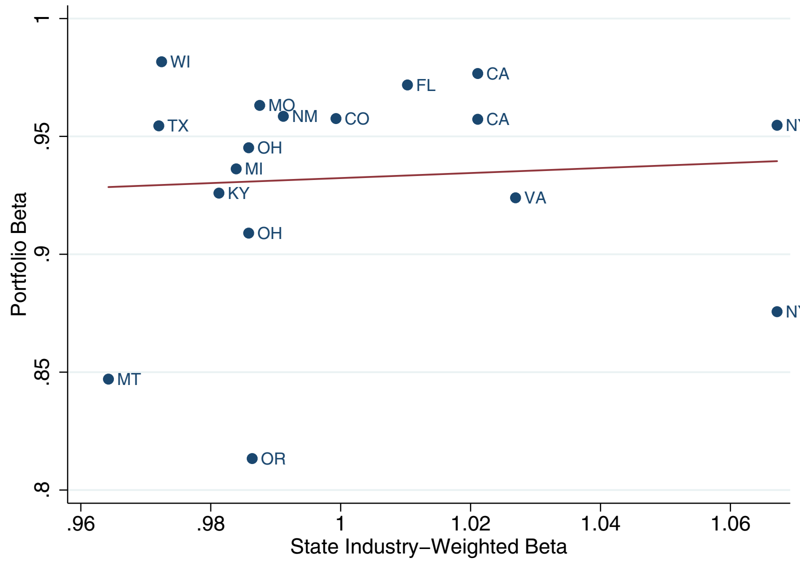
Note: This figure provides a visual display of the baseline measured 'excess' or 'windfall' returns. As described in the text, the baseline windfall return is equal to the realized return minus the baseline plan-specific benchmark return. This benchmark is equal to the weight the plan would have earned had it maintained its asset allocation and invested in national benchmarks. A detailed description of thes 'excess return' is available in the text. The mean of the 'excess return' distribution is 0.34 (measured in hundreds of basis points) and the standard deviation is 3.39.

Figure 6: Quintile of Excess Funds by Year



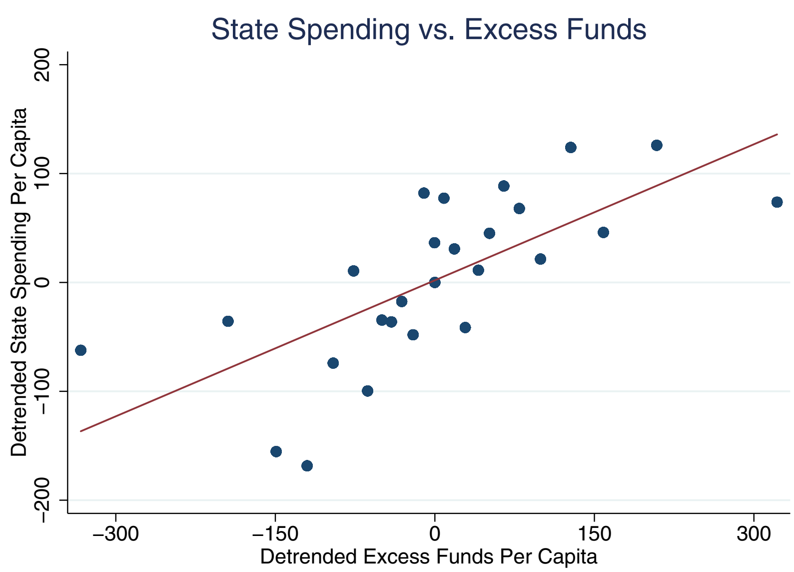
Note: These figures provide a visual display of the baseline excess funds variable over three consecutive years in the middle of the sample. The shading of states represents the quintile of the state's excess funds variable for that year, with lighter to darker shading representing lower to higher quintiles. The figures demonstrate that relative windfalls have no geographic or time-series patterns. Details on the construction of the excess funds series are available in the text. Oregon and New Jersey are omitted for the reasons discussed in the text.

Figure 7: Endogeneity within Asset Classes



Note: This figure graphs the relationship between the value-weighted beta of each plan's 13F domestic equity portfolio against the industry-weighted beta of the administering state. It demonstrates that there is little relationship between the risk characteristic of state income and the riskiness of the stocks chosen within the class of domestic equities. All data are from the fourth quarter 1999 filing. Data on asset-level betas were taken from the CRSP year-end assignments. Data on state industry composition for that year was taken from the BEA. Data on Industry level betas are from Damodoran (1996) http://www.wiley.com/college/damodaran/betas.html.

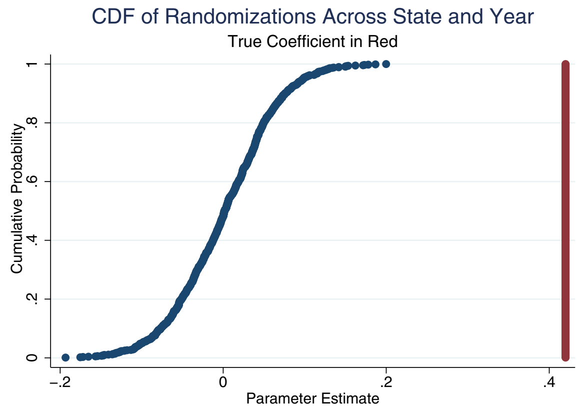
Figure 8: The First-Stage Relationship



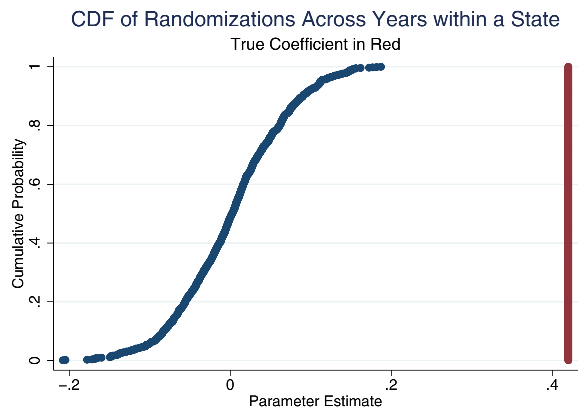
Note: This figure provides a graphical display of the first-stage relationship. The sample is split in to 25 equal-sized bins of the de-trended excess funds variable. For each bin, the population-weighted average de-trended excess funds and spending are plotted above. Both variables are in 2009 dollars per capita.

Figure 9: Placebo Tests of the First-Stage

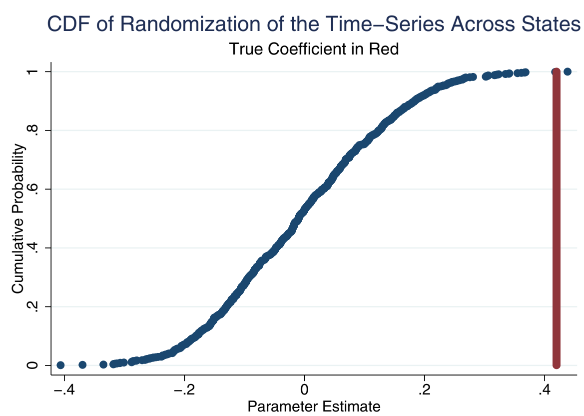
Panel A



Panel B

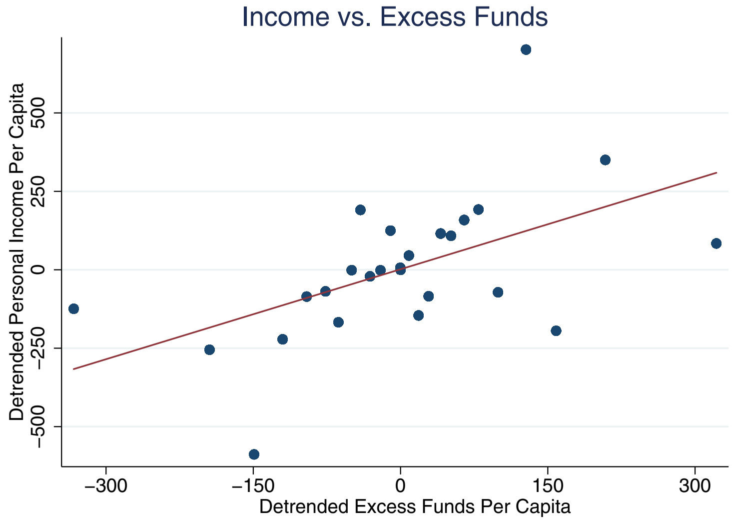


Panel C



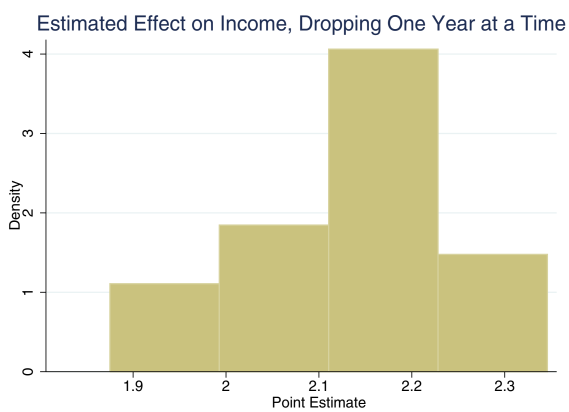
Note: These figures display the CDF from one-thousand Monte Carlo simulations. In Panel A, the excess funds series is randomly reassigned across all state and year observations. In Panel B, the excess funds series for each state is randomly reassigned across years in that state. In Panel C, the entire excess funds time-series is randomly reassigned across states with the time-series structure preserved. In each simulation, a coefficient is estimated for the regression of state government spending on the randomized excess funds variable, and state and year fixed effects. The regressions are weighted by population. In every panel, the coefficient estimated without randomization is displayed as vertical line in red. None of the randomized regressions in Panels A and B produce a point estimate equal to or larger than the true estimate, and only 1 randomization produced such a coefficient in Panel C.

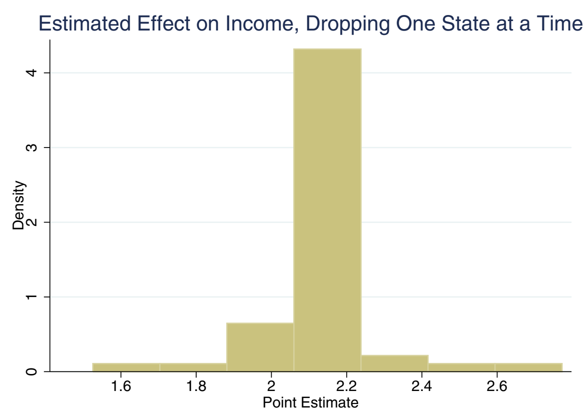
Figure 10: Graphical Display of the Reduced Form Income Effect



Note: This figure provides a graphical display of the reduced form relationship between de-trended income and de-trended pension windfalls. The sample is split in to 25 equal-sized bins of the de-trended excess funds variable. For each bin, the population-weighted average de-trended excess funds and income are plotted above. Both variables are in 2009 dollars per capita.

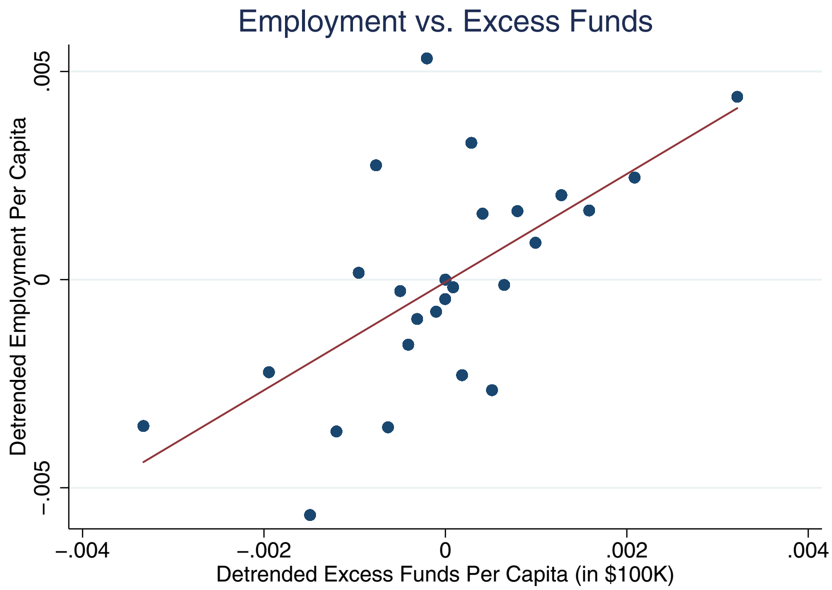
Figure 11: Robustness to Dropping States, Years





Note: The figures plot the histogram of estimated second-stage income coefficients when dropping one year (Panel A) and one state (Panel B) at a time. The coefficients are estimated using the baseline specification, which is in real-per capita levels and controls for state and year fixed effects. This specification uses the baseline instrument construction and is weighted by population The range of estimates in Panel A runs from 1.89 to 2.35. The range of coefficients in Panel B runs from 1.53 (when dropping CA) to 2.77 (when dropping OH).

Figure 12: Graphical Display of the Reduced Form Employment Effect



Note: This figure provides a graphical display of the reduced form relationship between de-trended employment and pension windfalls. The sample is split in to 25 equal-sized bins of the de-trended excess funds variable. For each bin, the population-weighted average de-trended excess funds and employment are plotted above. The excess funds variable is in 2009 dollars per capita. Employment is measured in employment per 100,000 residents.

Table 1: Survey on In-State Investment Bias

Survey Question: "What percent of the portfolio is *targeted* in-state?"
(Highlights in the original)

Year	1991	1993	1995	1997
Mean Response	0.3%	0.24%	0.27%	0.48%
% Answering Zero	95+	95+	95+	93

Source: Author's calculation from the PENDAT datasets

Table 2: Evidence on In-State Bias from 13F Filings

Percent of Portfolio In-State – Percent of SP500 In-State

Alaska Retirement Board	0.00%	New York State Retirement	−0.05%
California State Teachers	1.77%	New York State Teachers	0.18%
California Public Employees	0.38%	Public Employees of Ohio	0.45%
Public Employees of Colorado	−0.05%	State Teachers of Ohio	0.11%
Florida Retirement System	0.42%	Oregon Public Employees	−0.02%
Teachers Retirement of Kentucky	−0.01%	Pennsylvania Public School	1.36%
Michigan Retirement System	1.83%	Employees of Texas	−0.79%
Missouri State Employees	0.18%	Teachers Retirement of Texas	−1.42%
Montana Board of Investment	0.00%	Virginia Retirement System	0.06%
New Mexico Educational Board	0.00%	Wisconsin Investment Board	1.79%

Note: This table displays the difference between the share of a plan's portfolio invested in-state and the share of the SP500 invested in that state. It shows little in-state bias in the domestic equity investments of these plans. The average portfolio is 0.31% overinvested in state, and this estimate is not statistically significant in the sample. Data on state's portfolios are from the SEC's Edgar database. For 16 of the plans, the filing used was for the fourth quarter of 1999. This filing was not available for the remaining plans. The dates for the other filings are 6/30/2006 for the Alaska Retirement Board, 6/30/2000 for the PA Public School System, 3/31/2007 for the ERS of Texas and 9/30/1999 for the NY State Teachers Plan. Data on the composition of the SP500 are from CRSP and matched to state headquarters data from Compustat.

Table 3: **Evidence on In-State Bias from Investment Returns**

Panel A	(1)	(2)	(3)	(4)	(5)
Instrument	0.03	0.00	0.02	0.02	0.11
	(.02)	(.01)	(.37)	(.05)	(.10)
Description	Industry-Weighted Portfolio	Head-Quarters Portfolio	Contemporary Income Growth	Contemporary Employment Growth	Average Return in Neighboring State
R-Squared	.86	.86	.86	.86	.87
Observations	1592	1592	1592	1592	1100

Panel B	(1)	(2)	(3)	(4)	(5)
Instrument	−0.04	0.24	0.00	−.01	0.07
	(.03)	(.15)	(.01)	(.17)	(.19)
Description	Corruption Index	Election Year	Public Sector Union Contributions	Democratic Governor	Independent Investment Council
R-Squared	.86	.86	.86	.86	.86
Observations	1592	1592	1592	1549	1549

Note: The dependent variable is the annual fiscal year return measured in hundreds of basis points. These regressions demonstrate that state-specific economic shocks and political factors have little influence on plan returns. The regressions include, but do not report, fiscal year fixed effects. The standard errors are clustered by state. The instrument in Column (1) of Panel A is the return of the Fama-French industry portfolios weighted by state-industrial shares. The instrument in Column (2) of Panel A is the return on the value-weighted portfolio of stocks in the CRSP dataset that are headquartered in-state. Column (5) in Panel A restricts the sample to plans with June fiscal years, and calculates the un-weighted average return of plans in the neighboring states. Hawaii and Alaska are also dropped from this regression. The instrument in Column (1) of Panel B is the Glaeser-Saks (2004) corruption index. The instrument in Column (3) of Panel B is the share of political contributions made by public sector unions averaged over all the years in which data were available from the National Institute on Money in State Politics. Full descriptions of the instruments are available in the text.

Table 4: **Selected Summary Statistics**

	Mean	Std. Dev	Minimum	Median	Maximum
Retirement System Variables					
System Assets (in millions)	$24,094	$31,813	$95	$11,405	$269,358
Annual Return	10.2%	7.94%	-16.9%	11.7%	31.6%
SP 500 Return	12.5%	14.8%	-26.6%	13.6%	48%
Share in Equities	51%	17%	0	55%	83%
Share in Fixed Income	37%	17%	4%	33%	100%
Share in Short Term Assets	3%	4%	0%	2%	27%
Employer Contributions Per Cap	$117	$140	$0	$83	$1,348
Employee Contributions Per Cap	$55	$49	$0	$44	$341
Instruments					
Excess Funds (Baseline)	$27	$222	-$1,549	$17	$1,167
Detrended	.	$149	-$1,208	-$1	$848
Excess Funds (Broad)	-$18	$310	-$2,274	$0	$1,549
Detrended	.	$197	-$1,696	$1	$1,238
Excess Funds (Allocation)	-$44	$177	-$1,107	-$22	$918
Detrended	.	$131	-$747	$2	$846
State Variables [Per Capita]					
Total State Spending	$5,083	$1,800	$2,535	$4,743	$17,610
De-trended	.	$341	-$1,210	-$5	$3,259
State and Local Spending	$7,663	$2,147	$4,131	$7,273	$19,088
De-trended	.	$399	-$1,986	$2	$3,010
Personal Income	$33,994	$6,000	$20,748	$33,914	$57,501
De-trended	.	$1,077	$3,832	$20	$7,317
Employment (CES)	46%	4%	34%	46%	58%
De-trended	.	1%	-8%	0%	8%

Notes: See text and data appendix for sources. All data is in 2009 dollars. De-trended variables are the residuals from a regression on state and year dummies, weighted by population. Summary statistics for the additional variables used in this paper are available upon request.

Table 5: First-Stage under Different Constructions of the Instrument

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	g_t	g_t	g_t	g_t	g_t	g_t
Excess Funds$_{t-1}$	0.43***	0.37***	0.31***	0.25***	0.58***	0.24**
	(.12)	(.13)	(.10)	(.08)	(.19)	(.10)
R-Squared	.96	.88	.96	.96	.96	.96
Obs	964	1027	964	964	613	964
Fixed Effects	State, Year	State, Year	State, Year	State, Year	State, Year	State, Year
Instrument	Baseline 1987-2008	Baseline 1984-2008	Across and Within Macro-Benchmark A	Across and Within Macro-Benchmark B	Baseline Refined Asset Classes	Just Allocation System-Benchmark A

Note: This table demonstrates the effect of pension windfalls on state government spending. The regressions control unreported state and year fixed effects and are weighted by population. Standard errors are clustered by state. The dependent variable is per-capita dollars of state government spending, and the independent variable is per-capita dollars of pension windfalls earned over the prior fiscal year. The coefficient of interest can be read as dollars of spending per dollar of windfall. The instruments used in columns (1), (2) and (5) contain only the variation in pension returns that stems from performance within asset classes, as explained in the text. The instruments used in columns (3) and (4) contain variation from both asset allocation and performance within asset classes. The instrument in column (6) contains only variation in returns due to asset allocation. The 'excess returns' used in Columns (1) and (6) are only weakly correlated (ρ=.14), as discussed above. Additional details on the construction of these variables are available in the text.

Table 6: A Falsification Test of the Baseline Instrument

	(1)	(2)	(3)	(4)
	g_t	g_t	g_t	g_t
Excess Funds$_{t+1}$	0.11		0.07	
	(.07)		(.08)	
Excess Funds$_{t+2}$		-0.00		0.03
		(.08)		(.07)
Excess Funds$_{t-1}$			0.44***	0.48***
			(.08)	(.08)
Observations	964	964	868	868

Note: This table demonstrates that future value of the windfall series do not effect spending. The regressions control for unreported state and year fixed effects. The standard errors are clustered by state. The dependent variable is de-trended spending and the explanatory variables are from the de-trended excess funds series. Details on the construction of the excess funds series can be found in the text.

Table 7: Robustness Tests of the First Stage across Specifications

Variable	(1) g_t	(2) g_t	(3) g_t	(4) g_t	(5) $ln(g_t)$	(6) Δg_t	(7) g_t
Excess Funds$_{t-1}$	0.28**	0.18***	0.27***	.21***			0.43***
	(.12)	(.06)	(.07)	(.07)			(.09)
$\frac{\text{Excess Funds}_{t-1}}{g_{t-2}}$					0.18**		
					(.08)		
Δ Excess Funds$_{t-1}$						0.14**	
						(.07)	
Specification	FMB Cross-Section	FMB Time-Series	State Trends	State Quadratics	Logs	First Differences	Controls
Observations	21 Years	48 States	964	964	964	915	845

Note: The dependent variable is real per-capita state spending. The coefficients in columns (1)-(4) and (6)-(7) can be interpreted as dollars of spending per dollar change in the independent variable. The coefficient in Column (5) represents the percent change in spending from a windfall equal to the size of state spending two years beforehand. Column (1) reports the Fama-MacBeth coefficient derived by running 21 cross-sectional regressions of spending on the lag of excess funds. Column (2) reports the Fama-MacBeth coefficient from running 48 state-by-state regressions of spending on an unreported year trend and the excess funds variable. Regressions (3) through (7) include state and year fixed effects in addition to the listed controls. The standard errors in these regressions are clustered by state and the regressions are weighted by population. Six outliers in the change in first-difference of excess funds series are dropped in specification (6). The controls included in Column (7) are a state-level leading indicator created by the Philadelphia Federal Reserve, the annual per-capita dollars received from the Tobacco Master Settlement, the annual per-capita dollars received by the state in Federal grants, the initial balance in the state's general and 'rainy day' funds, and the initial size of the state's measured pension assets.

Table 8: The Timing of Windfall Spending, Positive and Negative Windfalls, and State Budget Cycles

	(1)	(2)	(3)	(4)	(5)
	g_t	g_t	g_t	g_t	g_t
Excess Funds$_{t-1}$	0.43***	0.43***	0.43***		0.49***
	(.12)	(.11)	(.11)		(.12)
Excess Funds$_{t-2}$		0.13	0.16*		
		(.08)	(.08)		
Excess Funds$_{t-3}$			.0		
			(.08)		
Excess Funds$_{t-1}$ (+)				0.55***	
				(.15)	
Excess Funds$_{t-1}$ (-)				.28**	
				(.13)	
Observations	964	915	846	964	741

Note: The dependent variable is the level per capita state spending. The regressions include state and year fixed effects and are weighted by population. Columns (1) through (3) explore the timing of the spending response from pension windfalls. Column (4) splits the pension windfalls in to positive and negative observations. Column (5) excludes biennially budgeting states in the off-year of their budget cycle. Coefficients can be read as dollars of spending per dollars of the independent variable. The independent variables are in levels of real per capita dollars and use the baseline instrument construction.

Table 9: The Spending Response by Level of Government

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Level of Spending	State	State and Local	Local	State to Local	Local	Fed to State	Federal	State
Excess Funds$_{t-1}$	0.43***	0.52**	0.25	0.15***	0.13	-.01	-.03	.046***
	(.12)	(.24)	(.17)	(.05)	(.14)	(.06)	(.21)	(.11)
State Grants$_t$					0.68***			
					(.19)			
Federal Grants$_t$								0.39***
								(.10)
Observations	964	876	876	964	964	900	964	890

Note: Data on state and local government spending are from the Census Bureau, and data on federal grants and spending are from the Consolidated Federal Funds reports. The variables are in real per-capita levels, and the regressions include state and year fixed effects and are weighted by population. The standard errors are clustered by state.

Table 10: The Spending Response by Functional Category

Composition	(1) Total	(2) Education	(3) Welfare	(4) Health & Hospitals	(5) Highways	(6) Corrections	(7) Retirement Benefits
Excess Funds_{t-1}	0.43***	0.14***	0.03	0.06***	0.04	0.02*	0.03**
	(.12)	(.04)	(.03)	(.02)	(.03)	(.01)	(.01)
Marginal Fraction		.43	.10	.18	.13	.07	.10
Average Fraction		.43	.25	.08	.12	.03	.07
R^2	.96	.88	.94	.82	.79	.80	.93
Observations	964	964	964	964	964	964	964

Note: The data are from the Census State Government Finance survey summary tabulations. The variables are in real per-capita levels, and the regressions include state and year fixed effects and are weighted by population. The standard errors are clustered by state.

Table 11: Effect of Windfall on State Tax Rates and Revenue

Panel A	(1) General Revenue	(2) Personal Income Tax Revenue	(3) Tax Burden	(4) Sales Tax Rate
Excess Funds_{t-1}	0.17**	0.17**	0.0003	0.0003
	(.08)	(.08)	(.0002)	(.0003)
R^2	.95	.95	.95	.87
Observations	964	964	964	885

Panel B	(1) Marginal Property Tax Rate	(2) Legislated Income Tax Revenue	(3) Legislated Total Revenue	(4) Cumulative Legislated $\text{Revenue}_{t,t+4}$
Excess Funds_{t-1}	-0.0006	-002	-0.03	-.03
	(.0006)	(.02)	(.03)	(.08)
R^2	.87	17	31	31
Observations	964	357	674	597

Note: The variables in Columns (1A), (2A), (2B) (3B) and (4B) are in real per-capita levels, and Columns (4A) and (1B) are in hundreds of basis points. All regressions include state and year fixed effects and are weighted by population. The standard errors are clustered by state. Data on tax revenue are from the Census State Government Finance survey. Data on the Tax Burden are from the National Tax Foundation, as is data for the Sales Tax Rate. Data for the marginal property tax rate is from the NBER TaxSim files. The data on legislated revenue changes is from the NASBO Fall Fiscal Year Surveys (www.nasbo.org) and only state-years with some recorded revenue change are used in the regression. Note that general revenue excludes revenue earned by state retirement system investments.

Table 12: **Instrumental Variable Estimation of the Effect of Government Spending on Income**

Panel A	(1)	(2)	(3)	(4)	(5)
	Income_t	Income_t	Income_t	Income_t	Income_t
Spending_t	2.12**	1.83**	2.08**	1.65	2.47*
	(.92)	(.77)	(.90)	(1.54)	(1.4)
First Stage F-Stat on Excluded	13.9	23.5	13.3	18.0	16.4
Specification	Baseline	Controls	Macro Benchmark A	Macro Benchmark B	Macro Benchmark B Outlier Trim
R^2 | Obs	.93 | 964	.93 | 845	.92 |964	.93 | 964	.92 | 946

Panel B	(1)	(2)	(3)	(4)	(5)
	Income_t	Income_t	$\ln(\text{Income}_t)$	ΔIncome_t	Income_t
Spending_t	2.39***	2.37**	0.38*	2.05**	2.39*
	(.72)	(1.11)	(.22)	(.87)	(1.26)
First Stage F-Stat on Excluded	21.5	10.8	4.7	4.6	15.1
Specification	Trim Largest 10% of Windfalls	State Trends Trends	ln(Spending) Scaled Excess Funds	Δ Spending ΔExcess Funds	Only Positive Excess Funds
R^2 | Obs	.92 | 868	.97 | 964	.94 |958	47 | 909	.92| 964

Note: The variables in Panel A, and in Columns (1), (2) and (5) in Panel B are measured in real per capita levels and dollars. All regressions include state and year fixed effects and are weighted by population. The standard errors are clustered by state. Column (2) adds as controls a leading state-level indicator created by the Philadelphia Federal Reserve, the annual per-capita dollars received from the Tobacco Master Settlement, the annual per-capita dollars received by the state in Federal grants, the initial balance in the state's general and 'rainy day' funds, and the initial size of the state's measured pension assets. Columns (3), (4) and (5) in Panel A use an alternate construction of the instrument that includes variation from differences in asset allocation. Details on its construction are contained in the text. Columns (3) and (4) in panel B drop the six largest values of the first difference of the instrument, as these extreme observations have a large effect on the results. The instrument in Column (3) is the excess funds variable scaled by the two-year lag of state spending. The endogenous explanatory variable and the dependent variable are the logs of per capita spending and income respectively. The coefficient represents an elasticity, which when interpreted at the sample means, yields a multiplier of 2.86. The instrument in Column (4) is the first difference of excess funds, the endogenous explanatory variable is the first difference of spending and the dependent variable is the first difference of income. Column (5) uses only the positive observations of the baseline windfall series.

Table 13: Robustness of Income Response to Using State and Local Spending

	(1)	(2)	(3)	(4)	(5)
	$Income_t$	$Income_t$	$Income_t$	$Income_t$	$Income_t$
State and Local $Spending_t$	1.84**	2.08**	2.28**	1.83**	2.24**
	(.74)	(.88)	(1.07)	(.72)	(1.03)
F-Stat on Excluded	4.9	6.8	8.7	4.7	9.3
Anderson-Rubin 90%-CI	[.2, 4.7]	[.6, 5.1]		[.2, 4.6]	
Specification	Baseline	Controls	5% Trim	Macro Benchmark A	Benchmark A 5% Trim
R^2 | Obs	.93 | 845	.93 | 845	.92 |845	.93 | 845	.92 |845

Note: The variables are measured in real per capita levels and dollars. The regressions include state and year fixed effects, are weighted by population, and the standard errors are clustered by state. Details on the construction of the controls are included in the text. The Census does not have information on local government spending in 2001 and 2003.

Table 14: Effect of Government Spending on the Components of Income

	(1)	(2)	(3)
	Retail Sales	Manufacturing Investment	$\frac{Capital\,Expenditure_t}{Assets_{t-1}}$
$Spending_t$	1.62	0.11	-.00002*
	(1.34)	(.08)	(.00003)
Observations	772	867	99,470

Note: The variables in specification (1) and (2) are measured in real per capita levels and dollars. All regressions include state and year fixed effects and standard errors clustered by state. Columns (1) and (2) are weighted by population and Column (3) is weighted by the lagged value of assets. Retail sales data are collected from the US Statistical Abstract. Data on manufacturing investment is from Wilson and Chirinko (2009). Firm level investment data is taken from Compustat. Details on the variables are available in the text and in the online appendix.

Table 15: Instrumental Variable Estimation of the Effect of Government Spending on Employment

Panel A	(1) Employment$_t$	(2) Employment$_t$	(3) Employment$_t$	(4) Employment$_t$
Spending$_t$	2.89***	2.44**	2.84***	2.58**
(in $100,000)	(1.0)	(.98)	(.72)	(1.0)
$ per Job	$34,585	$41,033	$35,148	$38,812
95%-CI	$10,353-$58,818	$8,529-$73,573	$17,687-$52,601	$8,542-$69,082
First Stage F-Stat on Excluded Instrument	13.9	14	13.9	14
Data Seasonally:	CES, Unadjusted	CES, Adjusted	LAU, Adjusted	QCEW Unadjusted
Obs	964	890	964	890

Panel B:	(1) Employment$_t$	(2) Employment$_t$	(3) Employment$_t$	(4) Employment$_t$	(5) Employment$_t$
Spending$_t$	1.87*	3.12***	3.58***	2.57*	2.03***
	(1.10)	(.97)	(1.12)	(1.32)	(.69)
$ per Job	$53,476	$32,051	$27,932	$38,911	$49,148
F-Stat on Excluded	22.6	23.55	10.8	4.7	11.6
Specification	Macro Benchmark B	Controls	State Trends	State and Local	Bartik Shocks
Obs	964	845	964	868	871

Note: The dependent variable is the level of employment per capita. Spending is measured in 2009 dollars in units of $100,000. The coefficients can be read as the number of jobs created per $100,000 of spending. The regressions include, but do not report, state and year fixed effects. The regressions are weighted by population and the standard errors are clustered by state. Confidence intervals for the reciprocal of the coefficient ($ per job) are calculated using the delta method. Panel A measures the number of jobs generated using different measures of employment. The first column of Panel B uses a different construction of the instrument that includes variation from differences in asset allocation. The additional controls used in Column (2) of Panel B are: a state-level leading indicator created by the Philadelphia Federal Reserve, the annual per-capita dollars received from the Tobacco Master Settlement, the annual per-capita dollars received by the state in Federal grants, the initial balance in the state's general and 'rainy day' funds, and the initial size of the state's measured pension assets. Column (4) of Panel B instruments for total state and local government spending. The weak-instrument robust confidence interval is in the text. Column (5) controls for predicted employment using the methods developed by Bartik (1991), applied at the three-digit NAICS level.

Table 16: The Mechanism Underlying the Employment Response

Panel A Dependent Variable	(1) Employment$_t$	(2) Unemployment$_t$	(3) Labor Force
Spending$_t$	2.85***	-.33	2.51***
(in \$100,000)	(1.0)		
(.39)	(.74)		
F-Stat on Excluded	13.9	13.9	13.9
Obs	964	964	964

Panel B Dependent Variable	(1) Private Employment$_t$	(2) State Employment$_t$	(3) State& Local Employment$_t$
Spending$_t$	2.15**	.16**	.45**
(in \$100,000)	(.94)	(.06)	(.14)
F-Stat on Excluded	14	13.9	14
Obs	890	875	875

Panel C Dependent Variable	(1) Gross Jobs Created$_t$	(2) Gross Jobs Lost$_t$	(3) Net Jobs Created$_t$
Spending$_t$	6.56*	2.87*	3.69*
(in \$100,000)	(3.9)	(1.72)	(2.21)
F-Stat on Excluded	13.4	13.4	13.4
Obs	742	742	742

Note: The dependent variables in Panels A, B, and C are in per capita levels. Spending is measured in units of \$100,000 per capita. All regressions include state and year fixed effects, are weighted by population, and use the baseline instrument construction. The standard errors are clustered by state. The dependent variables in Panel A are from the BLS Local Area Unemployment Statistics (LAU). The dependent variables in Panel B are from the BLS Current Employment Statistics (CES), and the dependent variables in Panel C are from the Business Employment Dynamics (BED) survey. To enhance the precision of the estimates in Panel C, the five largest values of the excess funds series in absolute value were dropped.

Table 17: Persistence of Income and Employment Effects

Panel A	(1)	(2)	(3)
	Income_t	$\sum\text{Income}_{t,\,t+1}$	$\sum\text{Income}_{t,\,t+2}$
$\sum\text{Spending}_{t,\,t+j}$	2.12^{**}	2.27^{**}	2.04^{*}
	(.82)	(1.06)	(1.06)
F-Stat on Excluded	13.9	11.8	6.9
Observations	964	915	867
Panel B	(1)	(2)	(3)
	Employment_t	$\sum\text{Employment}_{t,\,t+1}$	$\sum\text{Employment}_{t,\,t+2}$
$\sum\text{Spending}_{t,\,t+j}$	2.89^{***}	3.39^{***}	3.87^{**}
	(1.03)	(1.22)	(1.60)
F-Stat on Excluded	13.9	11.8	6.9
Observations	964	915	867

Note: This table explores the persistence of spending effects. If spending has long lasting effects, then the point estimates in the columns should be rising. In Columns (2) and (3) the dependent variable is the sum of income or employment per capita from the date of the spending shock through year t+1 and t+2, respectively. The endogenous explanatory variable is the sum of spending through the same horizon. The regressions are in real levels per capita and include state and year fixed effects. They are weighted by population and the standard errors are clustered by state. The instrument is the baseline excess funds measure, which is described in detail in the text.

Table 18: The Effect of Economic Slack on the 'Multiplier'

	(1A) Income$_t$	(1B) Income$_t$	(2A) Income$_t$	(2B) Income$_t$	(3A) Income$_t$	(3B) Income$_t$	(4A) Income$_t$	(4B) Income$_t$
Spending$_t$	3.00***	1.40	3.53***	1.61	2.25***	1.64	2.64***	1.62
	(1.30)	(1.41)	(1.04)	(1.76)	(.80)	(1.81)	(.79)	(2.02)
Sample	$e_{t-1} < \bar{e}_i$	$e_{t-1} > \bar{e}_i$	$lf_{t-1} < lf_{t-2}$	$lf_{t-1} > lf_{t-2}$	$e_{t-1} < .455$	$e_{t-1} > .455$	$lf_{t-1} < .5$	$lf_{t-1} > .5$
1^{st} Stage Coefficient	0.44***	0.28***	0.45***	0.34***	0.61***	0.24***	.51***	0.25***
F-Stat	11.9	11.1	11.9	5.5	20.8	9.2	17.0	7.9
Observations	411	553	349	615	426	525	370	590

Note: This table explores the impact of labor force slack on the income effect of government spending. The A and B specifications represent the estimation of the standard IV specification on two halves of the samples. The division in Columns (1A) and (1B) splits the sample into years in which the prior year had employment below the state's mean and employment above the state's mean, respectively. Columns (2A) and (2B) split the sample into years in which the prior year saw falling labor force participation and rising labor force participation, respectively. These divisions split the sample across years within states. Columns (3A) and (3B) split the sample based on whether employment in the previous year was below 45.5% or exceeded 45.5%. Columns (4A) and (4B) split the sample based on whether labor force participation in the previous year was below or exceed 50%. These divisions were chosen to maintain a sufficiently powerful first-stage despite the division. The standard specification is in real per capita levels, includes state and year fixed effects, and is weighted by population. The instrument is the baseline one described in the text. The standard errors are clustered by state.

Table 19: **Different Spending Effects in the Traded and Non-Traded Sectors**

	(1) Traded Income	(2) Traded Employment [Per $100K]	(3) Traded Avg. Wage	(4) Non-Traded Income	(5) Non-Traded Employment [Per $100K]	(6) Not Traded Avg. Wage
Spending ($)	-.32	.42	-0.57	2.35***	1.71***	2.98**
	(.51)	(.34)	(2.51)	(.71)	(.59)	(1.49)
Observations	890	846	846	890	846	846
F-Stat	14	13.7	13.7	14	13.7	13.7

Tracking the Changes ($1 in Spending Generates):

$2.98 Wage Increase * 29% Non-Traded Employment =	$0.86
.000017 Additional Non-Traded Jobs *$34,926 Average Non-Traded Wage =	$0.60
Direct Government Sector Income	$0.33
Total Measured Change	$1.79

Note: This table compares the effect of government spending across the traded and non-traded sectors. Following Rodrick (2009), I classify agriculture, mining, manufacturing and wholesale as traded industries. I classify retail, transportation, information, finance and insurance, real estate, all services, health care, recreation and the government sector as non-traded. Data on employment and average wages by industry come from the QCEW. Unlike the income measure, they do not contain the government sector. Wages are measured annually. All regressions include state and year fixed effects, use the baseline instrument, and are weighted by population. The standard errors are clustered by state.

Table 20: Spillovers, Migration and Housing Prices

	(1) In-State Counties	(2) Out-of-State Counties with Greater than 10% Working in State	(3) Out-of-State Counties Less than 10% Working in State	(4) Immigrants [Per Cap]	(5) Housing Prices [Percent Change]
Spending_t	2.51**	2.42**	-0.17		
	(1.12)	(1.08)	(0.27)		
Spending_t				0.51	
[$100K]				(.46)	
$\Delta\text{Spending}_t$					.01*
					(.00007)
Obs	61,401	66,587	694,607	928	792

Note: Columns (1), (2), and (3) explore the effect of government spending on county-level personal income using data from the BEA. The independent variable in these columns is the level of real per capita state spending. The dependent variable in Column (1) is the level of real per capita income for in-state counties. The dependent variable in Column (2) is the level of real per capita income for out-of-state counties in which greater than 10% of the workforce works in the treated state. The dependent variable in Column (3) is the level of real per capita income for out-of-state counties in which less than 10% of the workforce works in the treated state. Both Columns (2) and (3) include only out-of-state counties in neighboring states. Data on employment by location are from the Census' 1990 Journey to Work dataset, and to avoid endogeneity concerns, the sample is restricted to the post-1990 period. These regressions include county and year fixed effects and are weighted by county-level population. Data on cross-state immigration for Column (4) comes from the IRS SOI tax stats. Immigration is measured using the number of new exemptions filed in-state that had previously filed in another state. State level housing prices in Column (5) are measured using the Federal Housing Finance Agency seasonally adjusted purchase-only index. Columns (4) and (5) include state and year fixed effects and are weighted by population. The standard errors in all columns are clustered by state.

Table 21: Suggestive Evidence on the Spending Multiplier, Holding the Windfall Effect Constant

	(1)
	State Income Response to Windfalls
State Spending Response to Windfalls	2.17^{***}
	(.64)
R^2	.29
Observations	48

Note: The dependent variable is the coefficient β^{Inc} estimated from the state-by-state regression $y_t = \alpha_0 + \alpha_y * Year + \beta^{Inc} * Excess\ Funds_{t-1}$. The explanatory variable is the coefficient β^{Spend} estimated from the state-by-state regression $g_t = \alpha_0 + \alpha_y * Year + \beta^{Spend} * Excess\ Funds_{t-1}$. The coefficient in the table demonstrates the effect that increasing spending from windfalls has on the reduced form impact of windfalls on income. The point estimate indicates that a $1 windfall will increase income by $2.17 more in a state that spends the entire windfall ($\beta^{Inc} = 1$) than in a state that spends none of it ($\beta^{Inc} = 0$). Robust standard errors are presented in parenthesis.

Table 22: Additional Suggestive Evidence on the Spending Multiplier, Direct Windfall Effect

	(1)	(2)
	Spending_t	Income_t
$\text{Excess Funds}_{t-1}$	0.58^{***}	1.29^{***}
	(.13)	(.47)
$\text{1[Strong BB]*Excess Funds}_{t-1}$	-0.43^{***}	-1.10^{*}
	(.16)	(.65)
Coefficient Ratio	2.55	
$\left[\frac{Income\ Interaction}{Speding\ Interaction}\right]$	(1.61)	
	P-Value : .11	

Note: This table provides suggestive evidence on the effect of spending on income, holding the size of the windfall constant. Column (1) demonstrates that spending in states with stronger balanced budget rules responds less to pension windfalls. Column (2) demonstrates that the reduced form income effect in these states is also smaller. The bottom panel compares the ratio of the coefficient to produce an estimate of the spending effect, holding constant the direct effect of the windfall. The implicit assumption is that, save for the spending channel, the direct effect of the windfall on income is the same across the two groups of states. The variables are measured in levels and real per capita dollars. The regressions include state and year fixed effects and are weighted by population. The standard errors are clustered by state. The regressions are estimated as a system. States are divided into weak- and strong-balanced budget rule categories based on the ACIR index (1987). Following Clemens and Miran (2010), I use an index cutoff of 7 to classify strong-rule states. Note that the dummy for strong balanced budget rule is subsumed by the state fixed effects.