Estimating Variation in Productivity Across State Medicaid Programs: Evidence from Dual-Eligibles *

Timothy Layton† Nicole Maestas‡ Daniel Prinz§ Mark Shepard¶ Boris Vabson‖

May 15, 2024

Abstract

Many social programs involve some level of local autonomy, with local governments making some program design decisions within a set of national guidelines, and with financing being shared between local and national governments. We study the extent to which this autonomy results in heterogeneity across U.S. states in the productivity of their primary social health insurance program: Medicaid. We do this for one of the most expensive groups of Medicaid beneficiaries—those also enrolled in Medicare, i.e. "dual-eligibles". Productivity is typically defined as the ratio of outputs to inputs. For duals in Medicaid, we define inputs as the fiscal cost of the Medicaid program. We define output as providing access to healthcare goods and services, specifically access to primary care, critical surgeries, and, most importantly, long-term services and supports not covered by the Medicare program. We leverage duals who move across states to estimate state effects on fiscal costs, showing significant variation across states in these costs (a ratio of 3:1 for the highest to lowest spending states). In future work, we will also study the extent to which outcomes differ across higher- and lower-spending states.

*The research reported herein was performed pursuant to grant RDR18000003 from the US Social Security Administration (SSA) funded as part of the Retirement and Disability Research Consortium. The opinions and conclusions expressed are solely those of the author(s) and do not represent the opinions or policy of SSA, any agency of the Federal Government, or NBER. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation or favoring by the United States Government or any agency thereof. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

†Harvard University and NBER. Email: layton@hcp.med.harvard.edu

‡Harvard University and NBER. Email: maestas@hcp.med.harvard.edu

§World Bank. Email: dprinz@worldbank.org

¶Harvard University and NBER. Email: ff

‖Harvard University. Email: vabson@hcp.med.harvard.edu

Social insurance programs account for an increasing share of current and future government spending for countries around the world. They also have a major impact on people's lives, affecting health, wealth, employment and labor supply, human capital investment, fertility, family structure, and overall levels of well-being. But while we know that the *existence* of these programs has important impacts, outside of the effects of overall levels of generosity, little is understood about how these programs are designed and the consequences of different design decisions.

A principal design decision for any social insurance program is the level of local autonomy. Some programs are designed and managed at the national level, while other programs allow local governmental bodies to have some say in program design and management. In some cases, local governments can only modify programs within narrow bounds. In other cases, national governments simply allocate blocks of money to local governments, earmarked for a particular priority, and allow the local government full freedom to allocate those resources as it sees fit. In the United States, many programs, such as Medicaid, unemployment insurance, and the Supplemental Nutrition Assistance Program (SNAP), involve some level of local control. Yet we know little about the extent to which local control results in variation across local governments in (1) inputs (i.e., spending), (2) outputs, and (3) productivity (outputs for unit of input).

In this paper, we investigate this question in the context of the Medicaid program in the United States. With 81 million enrollees, the Medicaid program is the largest social insurance program in the United States in terms of beneficiaries. It also makes up a significant portion of state (29%) and federal (7%) budgets. Importantly, it involves significant local autonomy, with state governments being responsible for many design decisions, including rates paid to healthcare providers, benefit limits such as caps on the number of prescriptions a beneficiary can fill or the number of days of home and community-based services the program will pay for, and administrative hassles on the patient and provider sides for accessing care. Finally, there is also high-quality data on inputs (program spending) and outputs (access to care and health outcomes) at an individual level for all states. These features make this program ideal for measuring and understanding variation in productivity of a key social insurance program across states.

We focus on a key group within the Medicaid program, those eligible for both Medicaid *and* Medicare, typically referred to as 'duals.' For this group, Medi*care* covers the majority of most typical healthcare costs, including hospitalizations, office visits, prescription drugs, etc. Medi*caid* covers cost-sharing remaining after Medicare coverage, typically around 20% of the cost of care, as well as a few critical services not covered by Medicare, including long-term services and supports and behavioral health services. While it may seem odd to start with a group for whom Medicaid only covers a portion of healthcare spending, this group is one of the most expensive groups in Medicaid: Despite having the majority of most typical healthcare services covered by Medicare and making up only 14% of Medicaid enrollees, this group accounts for a full 30% of Medicaid spending. Further, for this group we can observe in Medi*care* administrative files high-quality, validated data on health outcomes and addresses that are consistently measured across states. Combining these outcomes with linked Medicaid administrative records, which we show to be capable of measuring higher-level outcomes such as spending and aggregate measures of utilization (though not health outcomes or detailed utilization measures) in most states, thus allows us to measure both inputs and outputs for duals by state. This group thus represents an ideal place to start when attempting to characterize variation in productivity of state Medicaid programs. Future work will focus on other Medicaid populations, where measurement is more difficult.

We start by leveraging the richness of our linked Medicaid-Medicare administrative data to provide a novel set of stylized facts about variation across state Medicaid programs for duals. After validating individual-level Medicaid administrative data on program spending using external spending records, we document enormous variation in per-enrollee program spending on duals across states. The highest-spending states have per enrollee spending levels that are 3-4 times the levels of the lowest-spending states. The bulk of spending for duals (and the variation) comes from long-term services and supports, with the most of that category of spending going to institutional long-term care in nursing homes.

We then investigate the extent to which variation in inputs (spending) across states is due to differences in the composition of Medicaid beneficiaries across states versus being due to 'causal' state effects. To do this, we follow Finkelstein et al. (2016) and leverage duals who move between states and thus between state Medicaid programs. We leverage continuous information on beneficiaries' zip codes of residence in the Medicare data to identify nearly 200,000 duals who move between states at some point during our sample period. We then implement the movers design of Finkelstein et al. (2016) to estimate that when a dual moves from a lower-spending state to a higher-spending state, their spending increases by around 60-70% of the observed cross-sectional spending gap between the states.

Event studies show that the spending of duals moving to higher spending states trends similarly to that of duals moving to lower spending states and to non-movers prior to the move. Patients who move to higher and lower spending states also look fairly similar, with prior spending having almost no predictive power for where a patient moves. At the time of the move, spending changes immediately, and the changes are generally persistent over time. The convergence is similar for moves from lower-spending states to higher-spending states and vice versa and for smaller and bigger (in terms of average spending in a patient's origin and destination state) moves. These results hold when we drop states that rely on private managed care plans to provide Medicaid benefits to duals, and where program spending consists of a capitation payment to those plans. They also hold when we look separately at under-65 duals (for whom eligibility rules are more standardized across states) versus over-65 duals (where eligibility rules can vary significantly across states).

We leverage the moving duals to estimate a two-way fixed effects model in order to recover causal state effects on spending. We find a tight linear relationship between the causal state effects on spending and the cross-sectional observational spending measures. The causal spending gap remains quite large: Our estimates indicate that the *causal* spending gap between the highest 10% of states and the lowest 10% of states remains large. We show that this causal spending gap is not unique to one type of spending but appears for all of the important types of spending for duals—Medicare cost-sharing, institutional long-term care, non-institutional long-term care, and other Medicaid-only services.

While our results clearly indicate important state effects on Medicaid spending, we are not interested in *all* types of state effects on spending. Indeed, it is useful to distinguish between state effects on spending due to program design decisions and other local factors such as provider practice patterns (Fisher et al. 2003a,b, Finkelstein et al. 2016, Cutler et al. 2019). When assessing differences in productivity across state Medicaid programs, we want to investigate variation in inputs driven by the former rather than the latter. In order to assess whether the state effects we've estimated are due to other local factors, we compare our state effect estimates to state effect estimates from a population not affected by Medicaid program design: non-dual Medicare beneficiaries. We then show that the state effects we estimate for this group are essentially orthogonal to our Medicaid state effects, implying that in order for practice patterns or other local factors to explain our Medicaid state effects, those patterns would have to be different from practice patterns for Medicare beneficiaries. Further, we show that state effects on the *Medicare*-portion of spending for our dual movers are also orthogonal to our estimates of state effects on the *Medicaid*-portion of spending for the same group. This further validates our conclusion that our main state effect estimates are capturing variation in spending due to program design rather than things like provider practice patterns that would be expected to influence Medicare spending as well.

In future work, we will compare our estimated state effects on Medicaid spending to other outcomes, including utilization of long-term services and supports. Preliminary results indicate little relationship between overall Medicaid spending effects and utilization.

1 Data Description

We use national administrative enrollment and claims data from Medicare and Medicaid.

Enrollment Data: The primary data source is linked administrative data from the Centers for Medicare and Medicaid Services (CMS) covering the Medicare and Medicaid programs during the 2008-2015 time period. The different Medicaid and Medicare administrative files are all linkable to one other, through a standardized beneficiary identifier.

The Medicaid and Medicare data includes separate enrollment files covering each of the programs. The observation level in all of the enrollment files is at a beneficiary-year level. The Medicaid as well as Medicare enrollment data track demographic information such as age, gender, and birth date. Importantly, the data also include the beneficiary's zip code, which comes from the Social Security Administration and is updated when the beneficiary movers. They also track each beneficiary's basis for eligibility at a given point in time, based on standardized eligibility codes. The eligibility information can be used to determine which Medicaid and Medicare enrollees are eligible due to disability vs. some other eligibility pathway. Birth date information, meanwhile, allow us to observe when individuals in the treatment group should move from Medicaid to Medicare. Finally, the enrollment files track actual enrollment status in both Medicaid and Medicare on a month-by-month basis, as well as whether an individual is simultaneously enrolled in both Medicaid and Medicare.

Utilization Data: The claims data track health care utilization across the full continuum of care types, including inpatient, outpatient, long term care, and prescription drugs. However, these data only track this utilization for a subset of all Medicare and Medicaid beneficiaries. For Medicare, the data tracks all utilization for those in Fee-for-Service (FFS), but does not track medical utilization for those in Medicare Advantage. For Medicaid, the data tracks utilization most reliably for those in fee-for-service Medicaid, and less reliably for those in Medicaid Managed Care.

The Medicaid and Medicare files contain standard claims data elements, including dates of service, unmasked National Provider IDs (NPIs) of prescribing and rendering providers, as well as diagnosis and procedure codes. These fields can be used to characterize and categorize the type of utilization provided.

For Medicare and Medicaid FFS, these data also track actual healthcare expenditures, in terms of actual amounts that are paid out to healthcare providers. The Medicaid Managed Care claims data, meanwhile, does not track actual amounts paid to providers.

Fiscal spending data: The Medicaid and Medicare claims files both track fiscal spending, in terms of the amount that the government spends on coverage for beneficiaries. The Medicare and Medicaid FFS claims data tracks fiscal spending directly, given that under the FFS programs, fiscal spending is equivalent to the amount that gets paid out to providers (since the government is directly on the hook for it). Meanwhile, for Medicaid Managed Care, fiscal spending comes in a different form: capitation payments made by the government, to private insurers. Fortunately, we can track these payment amounts in Managed Care encounter data that are included as part of the broader Medicaid claims dataset. Specifically, the data tracks capitation payment totals paid out to Medicaid Managed Care plans, at a beneficiary-month level.

1.1 MAX Data Quality

We start by investigating the quality of the Medicaid administrative data. Medicaid MAX data has long been suspected of having major quality issues, due to variation across states in reporting practices and issues surrounding the reporting of encounters by Medicaid managed care plans to states and then on to the federal government. While many variables are difficult or impossible to validate using external data (as the MAX data are the only source of those outcomes), our primary outcome—total fiscal spending—can be validated using external CMS data contained in CMS-64 reports. These reports include Medicaid program spending by type for a large variety of spending types. They are highly reliable, as these reports are used to determine federal matching (FMAP) payments to state Medicaid programs, and thus the federal government has a strong incentive to ensure that the spending numbers are accurate, as higher numbers equate to larger financial flows from the federal government.

We aggregate the types of program spending in the CMS-64 reports that appear in the claims (including direct payments for healthcare goods and services as well as capitation payments to managed care plans, but excluding administrative expenses) at the state-by-year level. We also construct analogous state-by-year spending measures using the Medicaid MAX files. We then compare the MAX spending measures to the CMS-64 spending measures to assess the completeness and reliability of each state's MAX data in each year of our sample.

We then use these comparisons to purge state-years that appear anomalous. To do so, for each state, we plot the time series of spending in the MAX data and in the CMS-64 reports. We first construct a ”liberal” sample where the two measures trend similarly and also are not too far apart from each other in any given year. We then construct a ”conservative” sample where the states must be almost identical in both trend and level.

Appendix Figure A1 presents the 64-MAX comparison of per enrollee spending for the year 2012. The color of the dots corresponds to the liberal, conservative, or other samples, as described above. We note two patterns from the figure. First, most states are quite close to the 45-degree line, suggesting that the MAX data on total fiscal spending is reasonably accurate, at least relative to the overall variation in spending across states. Second, when restricting to our liberal and conservative samples, the states for which the MAX data is clearly wildly off are removed, and the remaining states appear to have reasonably accurate data.

Appendix Figure A2 takes a different approach, showing the ratio of MAX to 64 spending over time for each of the samples—all states, liberal states only, conservative states only. Here, it is again straightforward to see (1) that the MAX data is reasonably good at capturing true spending trends over time and (2) the liberal and conservative restrictions do indeed help improve the overall accuracy of the MAX data.

While the validation exercise does not correspond exactly to the sample we use in the paper (we focus on duals, while the validation exercise was for all Medicaid enrollees) and we will often look at specific types of spending rather than overall spending, we believe the exercise does improve confidence that our estimation method will pick up real signal rather than just noise from differential reporting or differential data quality issues.

2 Descriptive Statistics

We now describe the raw variation in Medicaid spending across states as well as presenting summary statistics on the components of Medicaid spending for duals. For these descriptive statistics, we focus on a sample of duals who are enrolled in both Medicaid and Medicare for at least 9 months of the year in 2012.[1] Panel A of Figure 1 presents per capita state Medicaid spending on duals for each state passing our data quality checks described in Section 1.1 (i.e., the 'liberal' data quality sample).

The difference in spending between the lowest spending states is large—The lowest spending states (California, Arizona, South Carolina) all spend around \$10 thousand per dual-eligible beneficiary per year. On the other hand, the highest spending states (North Dakota, Connecticut) spend around \$40 thousand per dual-eligible beneficiary per year, around 4-times as much. Panel B removes states with moderate and high levels of enrollment of duals in private managed care plans to show that the high levels of spending variation across states remain, even when focusing only on states without managed care. For some analyses below, we will need to remove managed care states, as not all relevant outcomes are observable for them. However, this figure shows that even in this more restricted group of states, we will be able to make inferences about drivers of variation in state Medicaid spending, as the high overall levels of spending variation remain.

Figure 2 shows how Medicaid spending for duals is divided between four broad categories: Institutional long-term care (LTC), non-institutional long-term services and supports (HCBS), Medicaid secondary payer (cases where Medicare is the primary payer for the claim, but Medicaid covers beneficiary cost-sharing), and other Medicaid primary payer. Panel A shows this breakdown for states passing data quality checks and without managed care (managed care makes this decomposition difficult in many states). Panel B shows the overall breakdown averaged across all states, as well as average spending breakdowns for the highest and lowest spending states. Clearly, much of the variation across states comes from variation in spending on LTC and HCBS. This is not surprising, as states have significant flexibility in designing their long-term care benefits.

Because Medicare covers many types of healthcare for duals, state Medicaid spending patterns for duals may differ greatly from spending patterns for non-duals. Figure 3 presents a scatterplot of state per enrollee

[1] For those enrolled for fewer than 12 months, we construct 'annualized' spending measures by dividing their total annual spending by the number of months they're enrolled and multiplying by 12.

Medicaid spending for duals versus per enrollee Medicaid spending for other large Medicaid eligibility categories (disabled adults, non-disabled adults), with dots for each group in different colors. It is clear that the amount a state spends on duals is tightly correlated with the amount they spend on other adults in Medicaid. Thus, even though duals are a special group, the insights from this group likely apply to other Medicaid eligibility groups as well.

2.1 Importance of State Versus Sub-state variation

Ultimately, our goal is to assess the extent to which variation in Medicaid spending across states is driven by state Medicaid program design versus other factors. As a first step toward this goal, we assess the extent to which states explain individual-level variation in (log) spending. We consider three sub-state geographic units - counties, Hospital Service Areas (HSAs), and Hospital Referral Regions (HRRs). To assess the extent to which states explain the overall variation in spending across these three units, we run simple Ordinary Least Squares (OLS) regressions at the individual level of spending on state dummies. We then re-run that regression three times, replacing the state dummies with county dummies, HSA dummies, and HRR dummies. We then compare the R-squareds of these regressions to assess the extent to which these three sub-state geographic units explain more of the variation in spending than the state dummies.

Table 1 reports the R-squareds from each of these regressions. States explain a full 10% of the variation in individual-level spending among duals across the entire Medicaid program. Given the large variation in Medicaid spending, this is a fairly astonishing level of performance for such a simple regression model. Indeed, counties, HSAs, and HRRs only peform slightly better, despite (1), in the case of counties and HSAs, being *much* smaller geographic units and (2), in the case of HSAs and HRRs, being specifically designed to capture geographic variation in healthcare utilization.

Clearly, states appear to matter greatly for explaining variation in Medicaid spending. This can also easily be seen in the map in Figure 4 that shows the geographic variation in Medicaid spending. While not all state borders are easy to discern, many are, again showing the importance for states in explaining the variation in Medicaid spending.

While these descriptive results suggest that state program design is important for determining Medicaid spending levels, they could also stem from differences in the composition of Medicaid enrollees across states. These compositional differences could arise due to differences in state-level Medicaid eligibility rules or due to natural variation in where different people live. Further, differences in healthcare provider practice patterns could also potentially explain these geographic differences. In the next section, we leverage a quasi-experimental framework for differentiating between actual state effects on spending versus compositional differences and then present evidence suggesting that those state effects are due to program design rather than geographic variation in practice patterns.

3 Estimating State Effects

So far, we have documented wide variation in per-enrollee Medicaid spending across states. However, the source of this variation remains unclear. Prior work focusing on geographic variation in spending in Medicare posits two potential sources of variation: differences in the composition of beneficiaries (in health status and preferences for medical care) and differences in physician and hospital practice patterns (Finkelstein et al. 2016, Cutler et al. 2019). As we've emphasized, however, variation in Medicaid spending can also stem from a third factor: variation across states in Medicaid program design. Indeed, our goal in this paper is to isolate this third factor.

In this section, we attempt to isolate the variation due to state Medicaid program design. We start by separating variation in Medicaid spending across states into 'person' and 'place' factors, using the migration of Medicaid beneficiaries across states and a two-way fixed effects (TWFE) design. Our approach follows Finkelstein et al. (2016), which in turns builds on work in labor economics that considers the role of firms in earnings inequality (Abowd et al. 1999, Card et al. 2013).[2] This decomposition will allow us to identify the portion of variation in Medicaid spending across states that is due to 'state effects'. Importantly, these state effects will include both the effects of Medicaid program design and the effects of variation in practice patterns. As we are interested in isolating the former, at the end of this section we will present additional analyses that establish the extent to which the state effects reflect program design versus practice patterns.

3.1 Empirical Framework

We start from a simple statistical model of Medicaid spending:

$$y_{ist} = \alpha_i + \gamma_s + \tau_t + x_{it}\beta + \epsilon_{ist} \tag{1}$$

where i indexes patients, s indexes states, t indexes years, y_{ist} is a measure of spending, the α_i are patient fixed effects, the γ_s are state fixed effects, the τ_t are year effects, and x_{it} is a vector of observable characteristics. The patient and state effects in this model can be separately identified if some of the patients move across states. We thus construct a sample of moving dual eligibles to estimate this model. Specifics regarding the construction of this sample are discussed in Section 3.2.

This model can also be represented as an event study that allows us to graphically display our results and assess the validity of the identifying assumptions. Let $o(i)$ denote patient i's origin state and $d(i)$ denote her destination state. (For patients who do not move, $o(i) = d(i)$.) Let $r(i,t)$ denote years relative to the year patient i moves (e.g. $r(i,t) = -2$ two years before patient i moves), which we will refer to simply as "relative years". To transform Equation (1) into an event study, we define $\delta_i = \bar{y}_{d(i)} - \bar{y}_{o(i)}$, the difference in average spending between patient i's origin and destination states. Combining the origin place effect $\gamma_{o(i)}$ with the patient effect α_i from Equation (1) into a single patient effect, $\mu_i = \alpha_i + \gamma_{o(i)}$, we can write down an event study as

$$y_{it} = \mu_i + \theta_{r(i,t)}\hat{\delta}_i + \tau_t + x_{it}\beta + \epsilon_{it} \tag{2}$$

where i indexes patients, t indexes years, y_{it} is a measure of spending, μ_i is a patient effect, $\hat{\delta}_i$ is the estimated difference between patient i's origin and destination states in average spending, the $\theta_{r(i,t)}$ are

[2] Other studies in health economics that use patient migration to separately estimate patient and place effects include Moura et al. (2019), Godøy and Huitfeldt (2020), and Salm and Wübker (2020) among others. There are also studies that exploit physician migration in order to separately identify physician and place effects, including Molitor (2018).

relative-year-specific coefficients, the τ_t are year effects, and x_{it} is a vector of observable characteristics. The $\theta_{r(,t)}$ coefficients show where the average spending of movers falls for each relative year compared to the average spending levels in their origin and destination states. For example, if on average movers' spending is the same as the average in their origin state, then $\theta_{r(i,t)} = 0$ and if on average it is the same as the average in their destination state, then $\theta_{r(i,t)} = 1$.

The model makes several important simplifying assumptions.[3] First, health shocks that coincide with the time of the move *and* are correlated with spending in the origin and destination would bias our estimates of state effects. One potential violation would be if the need for a nursing home triggers a move by a dual, *and* makes them more likely to move to a state with higher nursing home use. We provide a few pieces of evidence that our analysis does not suffer from this type of problem. First, we find no evidence in our event study analysis of pre-move trends in spending, showing that movers with different origin-destination differences in spending are on similar spending trends (see Figure 6). Second, we investigate whether there are spikes to spending at the time of the move. Finally, for our most powerful test, we examine how the pre-move utilization pattern of movers differs from matched non-movers depending on the type of move they make. Figure A5 shows the relationship between the "size" of moves and the difference in the pre-move utilization of movers and matched non-movers during the pre-move period. The slope of 0.03 here indicates that there is little selection on pre-move spending levels and any differences are absorbed by the patient fixed effects α_i.

Second, we assume that the patient effects α_i and state effects γ_s are additively separable. This rules out the possibility that different types of patients would behave differently within a state. For example, the causal effect of a place cannot be larger for high-spending patients. We test this assumption by including specifications with spending in levels (where we assume place and person effects are additive) and specifications with spending in logs (where we assume place and person effects are multiplicative).

Third, because state effects are identified only by movers, for state effects to have a broader interpretation, we need to assume that state effects for movers and non-movers are similar. If state effects are different for movers, our estimated state effects would not be valid for entire population. As one piece of evidence that movers and non-movers may be relatively similar, Table 2 shows summary statistics for the two groups. The table shows that the movers and non-movers look fairly similar on age, gender, though they do differ on healthcare spending. Medicare spending is somewhat similar, but Medicaid spending differs, with movers spending less than non-movers.

Fourth, this model does not allow past spending y_{it} to influence patient effects α_i, i.e. habit formation. Habit formation would mean that the estimated patient effect α_i could be partly driven by past state effects γ_j. We test for this possibility via the event studies of the effects of moves. Habit formation would appear as a gradual convergence of beneficiary spending to the spending level of their destination state. If effects appear immediately, on the other hand, habit formation is relatively unimportant. Ultimately, we find the latter.

[3] For a more detailed discussion, see Finkelstein et al. (2016).

3.2 Sample Definition

For this exercise we use the Medicare files from 2007 to 2015. Because we are interested in identifying movers, we drop beneficiaries only observed for one year. We define a patient's origin state as the first state that they appear in.[4] We define non-movers as beneficiaries who are always in the same state and movers as beneficiaries who are in two states during the timeframe. For our event study, we use only movers and drop non-movers. We define as the move year (relative year 0) as the first year that a patient is not in their origin state (the first state they appear in) but in their destination state (the second state they appear in). We limit to beneficiaries with no additional cross-state moves in the five years after the initial move, including back to the original source state or to a third state.

Further, we limit to beneficiaries who are enrolled in Medicare for all twelve months in relative years -2, -1, and 0 and in Medicaid in relative years -1 and 1. (This allows for some time to switch Medicaid enrollment from the origin state to the destination state.) We also drop years with incomplete (fewer than 12 months) Medicare or Medicaid enrollment. For analyses using outcomes from Medicare data, we construct a separate sample of beneficiaries never enrolled in Medicare Advantage. Finally, we require that movers have at least 75% of their spending in their source state pre-move and in their destination state post-move. This should eliminate beneficiaries with multiple homes, such as "snow birds."

3.3 Event Study Results

We start by establishing that patients categorized as movers do in fact move. Appendix Figure A3 shows the share of claims in the source and destination states in each relative year before and after the move. For each claim, state is defined as the state Medicaid agency that submitted data for that claim. It shows that prior to the move almost no claims are in the destination state. Then at the time of the move, the share of claims jumps to close to 100% and stays at that level in subsequent years. This suggests that we are indeed correctly identifying patients who change their state of residence. This is not surprising given the sample restriction that at least 75% of claims be in the destination state in post-move years.

Next we summarize the distribution of the magnitude of moves in Appendix Figure A6. This figure shows the distribution of origin-destination differences in total spending ($\hat{\delta}_i$). It suggests that the distribution is approximately symmetric with a mean close to zero, which means that there is about the same number of moves from lower-spending states to higher-spending states and vice versa.

To start understanding how spending changes when patients move, Figure 5 presents a binned scatter plot of the average change in spending across ventiles (twenty equal-sized bins) of the origin-destination difference in spending. In addition, the figure shows the line of best fit from a simple OLS regression, as well as the average change for a matched sample of non-movers. The slope of the line is 0.60 which suggests that a move from a state to another state with 1 log points higher (lower) spending is associated on average with a 0.6 log points increase (decrease) in spending. The relationship is close to linear across the distribution of origin-destination differences, suggesting that the effects of moves "up" (to a higher spending

[4] We use the Medi*care* data to identify state of residence, not the Medi*caid* data. The Medicare data continuously measures residence using data from the Social Security Administration. When duals move across states, they typically experience a gap in their Medicaid coverage, as they need to sign up for coverage in their new state (it is not automatic). During that gap, we do not observe state of residence in the Medicaid data, but we do in the Medicare data.

state) and moves "down" are fairly symmetric. We can interpret this as states effects explaining 60% of the cross-state variation, while patient effects explain 40%.

Figure 6 shows our main event study results for Medicaid spending. It suggests that pre-trends before moving are negligible. The coefficients of zero in these years suggest that on average movers' spending is the same as the average spending in their origin state. Then in the year of the move, spending converges discontinuously. The coefficient of 0.56 implies that the state share of spending differences is 56%, while the patient share is 44%.

Table 3 presents results from a set of regressions where we pool pre- and post-years into single pre- and post-periods in order to estimate a single θ coefficient that represents the average convergence in total spending pre- vs. post-move. The interpretation of this coefficient is as follows: A \$1 increase in the cross-sectional difference in average spending between the origin and destination states will result in a \$$\theta$ increase in a mover's spending change upon moving. We run several different versions of this regression to show the robustness of our estimates of θ. Specifically, we test robustness to further limiting the sample to state-years with even higher-quality data (our "conservative" sample), limiting the sample to states that do not use managed care (where spending will be determined by premium payments to Medicaid Managed Care (MMC) plans rather than by utilization), considering spending in levels versus logs, etc.

In our preferred specification (liberal sample and MMC strict), the coefficient is between 0.6 and 0.7, implying that state effects explain around 60-70% of the variation in total spending. This is a really significant share and implies that states really matter greatly for duals' Medicaid spending. The coefficient is higher when we include all states, possibly reflecting poor data quality. It is low for the MMC loose sample, where we keep states with high MMC share but drop people in MMC, suggesting that the movers analysis may not work well in those types of states where the remaining enrollees (not in MMC) are not representative of the state enrollment pool.

3.4 State Effect Estimates

We now turn to estimating our main model specified in Equation (1). The results presented in the previous section suggest that patient migration can be used to separately identify state and patient effects and that our key assumptions are satisfied.

Our estimates of state effects on log total Medicaid spending are found in Figure 7. The left panel shows state effects for the full liberal sample, while the right panel shows state effects for the sample of states that do not use managed care. The figures show that the variation in causal state effects on spending remains large, with the difference in spending between the states with the highest and lowest spending effects being around 2.5 log points. This gap in spending effects remains when restricting to states without managed care, showing that the large differences in spending effects are not driven solely by differences in premium payments made to managed care plans.

The states that have the largest positive spending effects (highest causal spenders) include Arizona, Minnesota, Utah, and Oregon. The states with the largest negative spending effects (lowest causal spenders) include Florida, South Carolina, Alabama, and West Virginia.

Table 4 summarizes our estimates of state effects. This table presents an explicit decomposition of the total variation in spending into the portion due to place (state) and patient heterogeneity. As in the prior analyses, this decomposition indicates that about 60% of the variation in spending comes from state effects, while the other 40% comes from person effects. This decomposition is relatively constant throughout the distribution of state effects, though it varies some, ranging from 60-70% at different parts in the distribution.

3.5 Effects of State Medicaid Program Design Versus Variation in Practice Patterns

Prior work exploring geographic variation in Medi-*care* utilization and spending points to variation in provider practice patterns as an important factor driving this variation (Cutler et al. 2019, Finkelstein et al. 2016, Cutler et al. 2013, Fisher et al. 2003a,b). We argue that in Medi-*caid* differences across states in the design of their Medicaid programs is likely to be the primary driving factor. However, practice patterns may still matter, and the state effects we have estimated will clearly include both.

When assessing the overall productivity differences of Medicaid across states, we need not distinguish between these two sources of variation in spending and utilization, as both contribute to the variation in inputs. However, variation caused by differences in program design is clearly more actionable than variation caused by differences in practice patterns. Further, our original intent was to try to answer the question of how the spending and outcomes of the Medicaid enrollees in one state (e.g., Texas) would change if that state adopted the Medicaid program of another state (e.g., New York). Clearly, to answer this question, we need to separate effects of program design and effects of variation in practice patterns.

While separating these two factors is difficult, we provide a test of whether practice patterns explain the variation in Medicaid spending we estimate. The test is simple - If practice patterns are responsible for the variation, then state effects on Medi-*care* spending should be highly correlated with state effects on Medi-*caid* spending. Medicare is administered and designed at the national level, leaving little or no room for state policies to cause variation in spending (especially state Medicaid policies). If, on the other hand, there is no relationship between state effects on Medicaid spending and state effects on Medicare spending, this will provide suggestive evidence that practice patterns do not explain our state effect estimates and instead state Medicaid program design is the factor driving the variation we estimate.

We implement this test in two ways. First, we estimate state effects on Medicare spending for non-duals (similar to the analysis in (Finkelstein et al. 2016)). Second, we estimate state effects on the Medicare-paid part of the spending of duals. In both cases, we continue to use the same movers design as before.

Figure 8 presents the results of these tests. In both panels, we plot the state effects on Medicaid spending versus the state effects on Medicare spending. The left panel shows Medicare state effects for non-duals, and the right panel shows Medicare state effects for the Medicare-paid portion of spending for our duals movers. In both cases, our estimates of state effects on Medi-*caid* spending are essentially orthogonal to the estimates of state effects on Medi-*care* spending, with slopes of 0.05 and -0.03. This suggests that the state effects we estimate are unlikely to be driven by more generaly variation in provider practice patterns, as this would likely result in a tight correlation between these sets of state effects.

4 Conclusion

Social insurance programs around the world exhibit varying levels of local autonomy in program design. We study one of the largest social insurance programs in the world that provides a great deal of local autonomy: Medicaid. We document significant variation across states in per enrollee spending among program beneficiaries also eligible for Medicare ("duals") who rely on this program for coverage of long-term services and supports and supplemental coverage for other services. We then leverage beneficiaries moving between states to show that much of this variation is driven by causal state effects on spending rather than differences in the composition of Medicaid enrollees across states.

These results indicate that state program design has important consequences for state spending in Medicaid. Some states spend much, much more than others. This is an important finding, as it suggests that the flexibility provided to states allows them to influence spending levels to a high degree. The next question is whether states that spend more also achieve better outcomes or whether they instead spend more with little to show for it. In future work, we will explore state effects on these outcomes in order to provide new evidence on heterogeneity in productivity across state Medicaid programs.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333. 7

Card, David, Jörg Heining, and Patrick Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 2013, *128* (3), 967–1015. 7

Cutler, David, Jonathan S. Skinner, Ariel Dora Stern, and David Wennberg, "Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending," *American Economic Journal: Economic Policy*, 2019, *11* (1), 192–221. 2, 7, 11

_ , Jonathan Skinner, Ariel Stern, and David Wennberg, "Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending," Working Paper 19320, National Bureau of Economic Research 2013. 11

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams, "Sources of Geographic Variation in Health Care: Evidence from Patient Migration," *Quarterly Journal of Economics*, 2016, *131* (4), 1681–1726. 2, 7, 8, 11

Fisher, Elliott S., David E. Wennberg, Therese A. Stukel, F. L. Lucas Daniel J. Gottlieb, and Etoile L. Pinder, "The Implications of Regional Variations in Medicare Spending: The Content, Quality, and Accessibility of Care. Part 1," *Annals of Internal Medicine*, 2003, *138*, 273–287. 2, 11

_ , _ , _ , _ , and _ , "The Implications of Regional Variations in Medicare Spending: The Content, Quality, and Accessibility of Care. Part 2," *Annals of Internal Medicine*, 2003, *138*, 288–298. 2, 11

Godøy, Anna and Ingrid Huitfeldt, "Regional Variation in Health Care Utilization and Mortality," *Journal of Health Economics*, 2020, *71* (1), 102254. 7

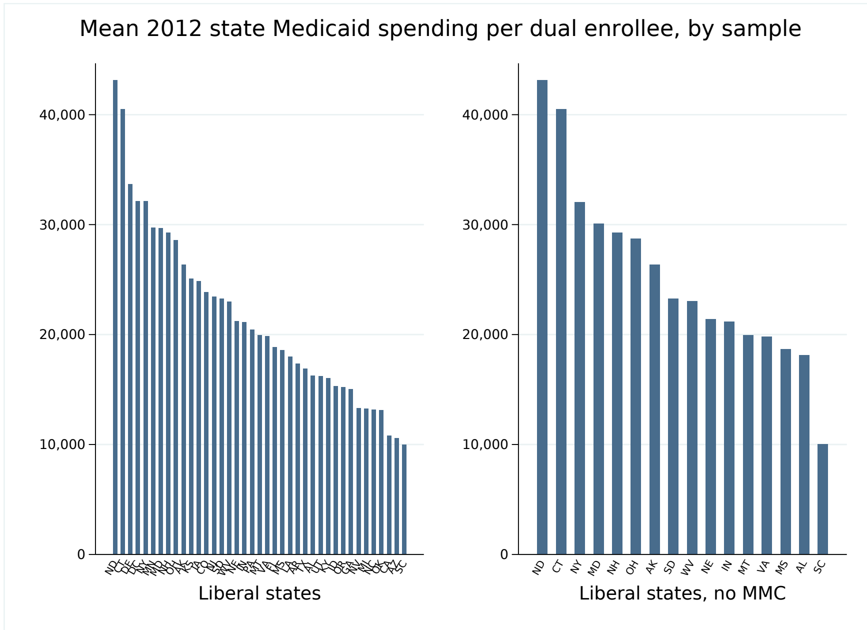
Molitor, David, "The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration," *American Economic Journal: Economic Policy*, 2018, *10* (1), 326–56. 7

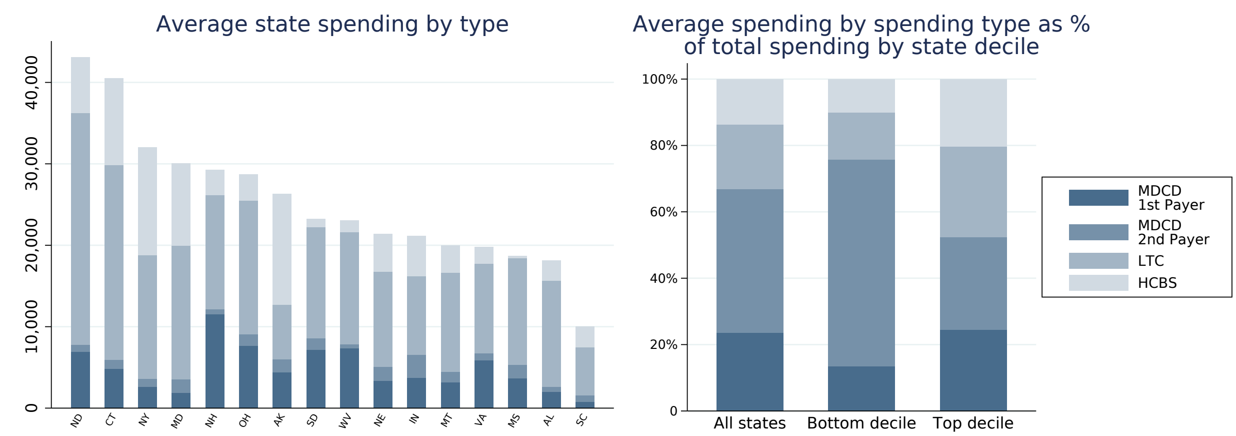
Moura, Ana, Martin Salm, Rudy Douven, and Minke Remmerswaal, "Causes of Regional Variation in Dutch Healthcare Expenditures: Evidence from Movers," *Health Economics*, 2019, *28* (9), 1088–1098. 7

Salm, Martin and Ansgar Wübker, "Sources of Regional Variation in Healthcare Utilization in Germany," *Journal of Health Economics*, 2020, *69* (1), 102271. 7

5 Figures

Figure 1: Variation across States in Per-Capita Medicaid Spending among Duals





(a) Liberal no MMC restriction: state Medicaid spending by category

(b) Liberal no MMC restriction: Medicaid spending by category and decile

Figure 2: Breakdown of Duals' Medicaid Spending

Figure 3: Correlations between Per Enrollee Medicaid Spending on Duals and Other Medicaid Eligibility Categories

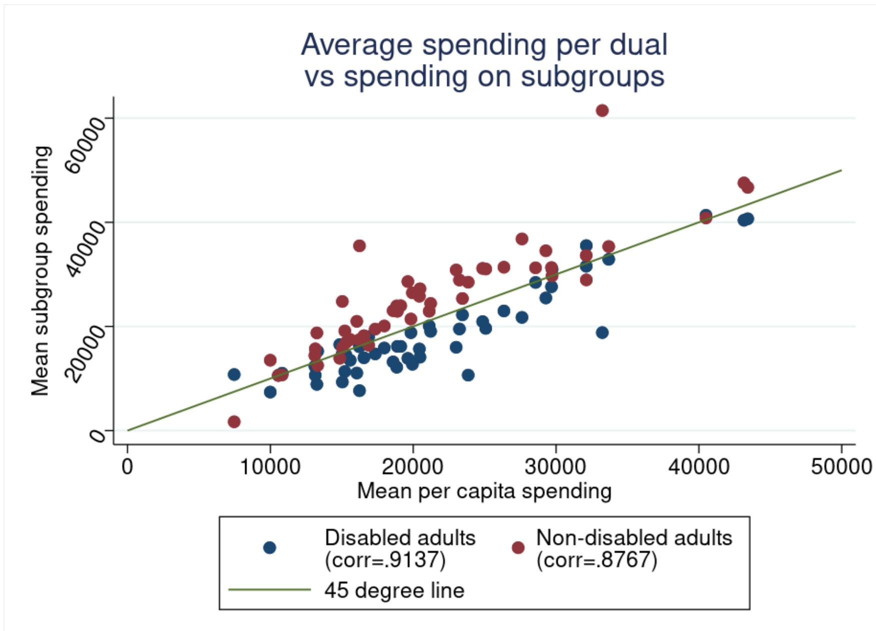


Figure 4: Map of per enrollee Medicaid spending among duals at the county level

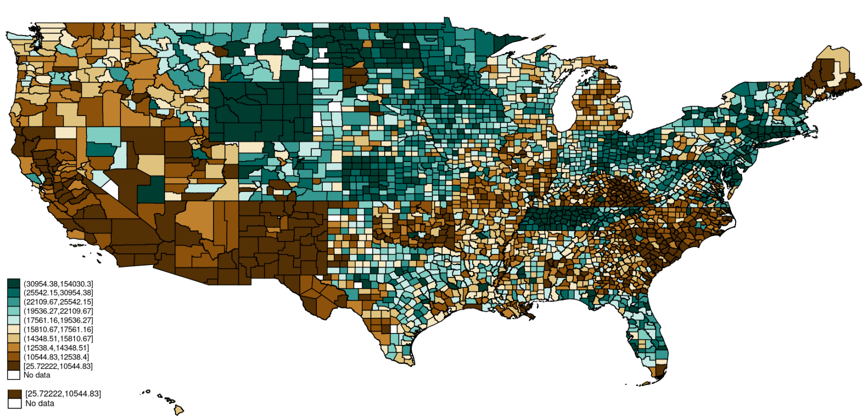


Figure 5: Binscatter of Pre- Vs. Post-Move Change in Spending Versus Difference between Origin and Destination State Average Spending

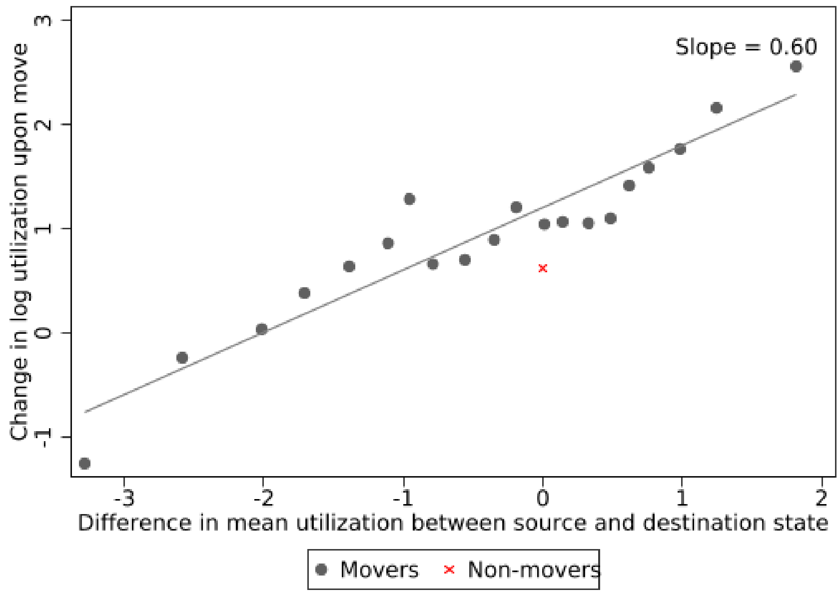
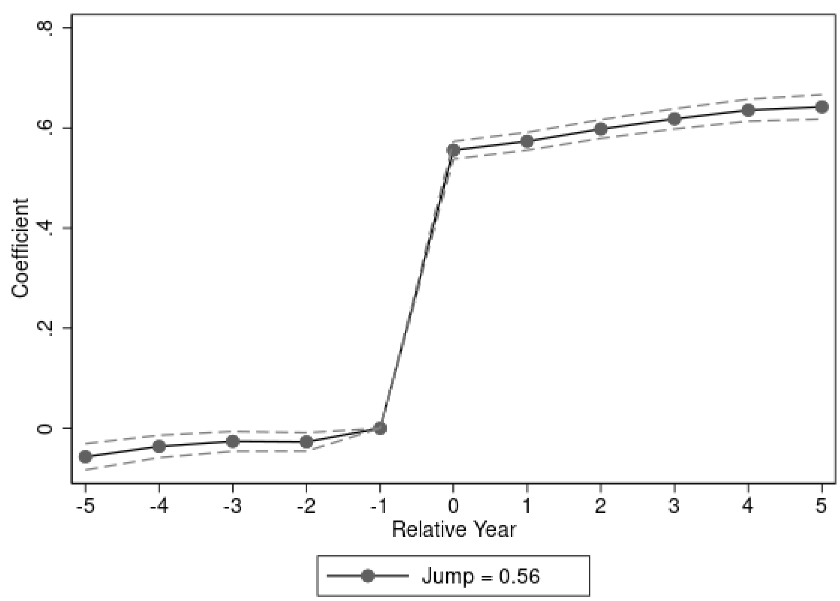
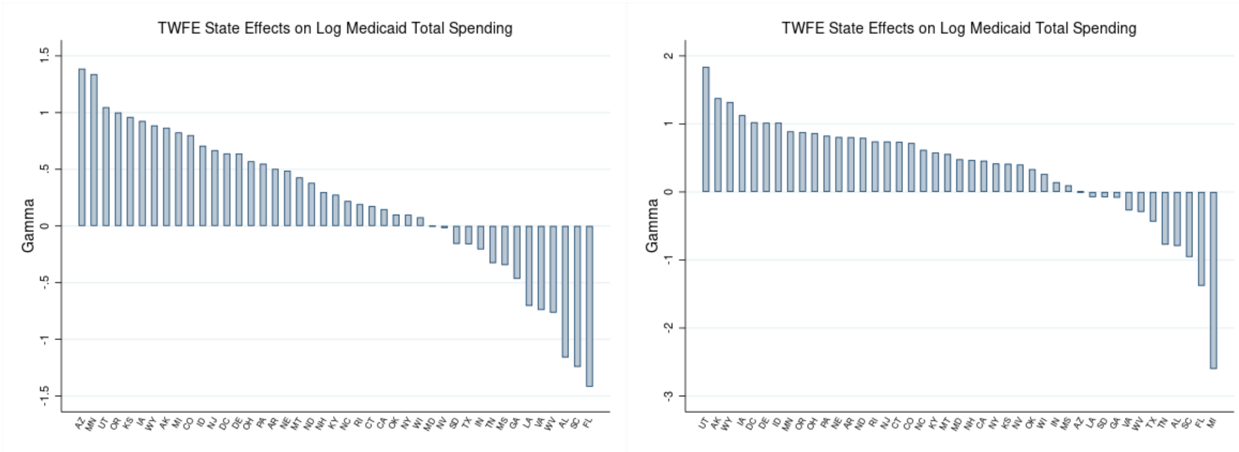


Figure 6: Event Study of Changes in Beneficiary Spending Relative to Origin-Destination Difference in Spending

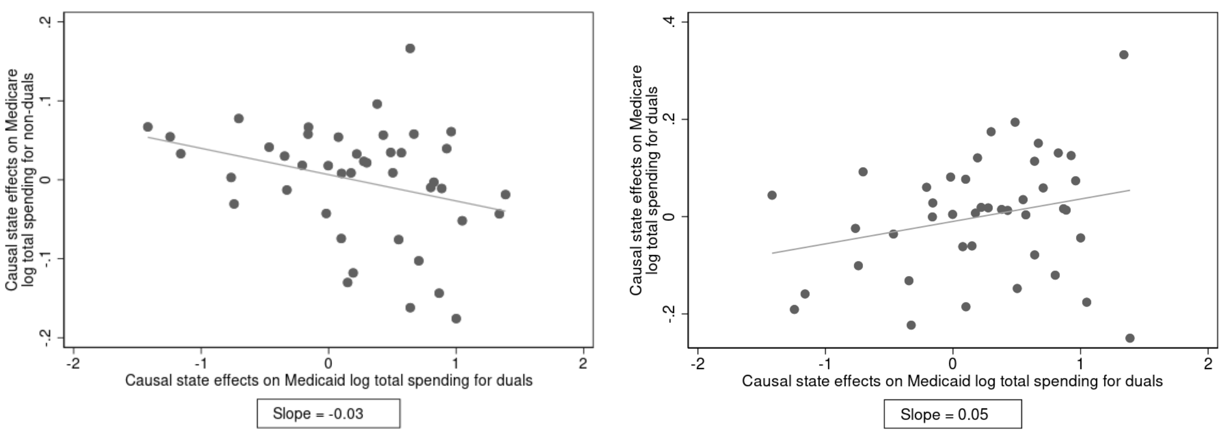




(a) Liberal restriction

(b) Liberal no MMC loose restriction

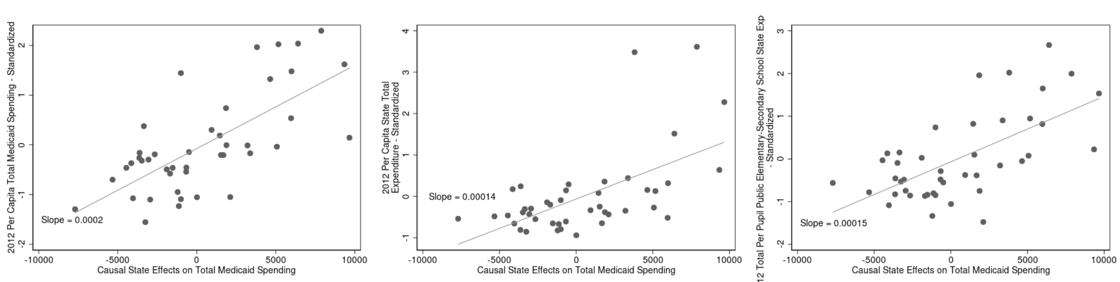
Figure 7: State Effects on Total Medicaid Spending



(a) Medicaid effect vs non-dual Medicare effect

(b) Medicaid effect vs dual Medicare effect

Figure 8: Medicaid Versus Medicare Spending Effects



(a) Total per enrollee Medicaid spending - 2012

(b) Total state fiscal spending per capita - 2012

(c) Primary + Secondary education spending per student - 2012

Figure 9: Spending Effects versus State Fiscal Practices

6 Tables

	Total Medicaid Spending	Log Total Medicaid Spending
State	0.0392	0.1089
County	0.0722	0.1323
HRR	0.0433	0.0966
HSA	0.0905	0.1277

Table 1: R-squareds from Regressions of Medicaid Spending on State, County, HSA, and HRR Fixed Effects

	Non-Mover Mean	Non-Mover SD	Mover Mean	Mover SD
Female	0.619	0.486	0.615	0.487
Age	63.519	16.635	62.071	17.711
Total Spending	25385.965	49278.685	18584.241	38500.350
Medicaid Spending	13386.941	41051.579	8521.830	27107.642
Medicare Spending	11999.024	23735.910	10062.411	24597.537
Medicaid Secondary Payer	858.541	11240.375	816.420	19053.640
Mdcd Primary LTSS	593.103	6056.363	1360.580	7027.412
Mdcd Primary LTC	6435.174	36176.743	3184.109	15878.416
Mdcd Primary Capitation	2787.944	9921.641	1770.813	6744.456
Mdcd Other Primary	3902.947	14209.453	2410.332	7699.907
Medicare Drug Spending	4975.479	8478.838	5023.211	8899.947
Medicare IP Spending	2864.119	11606.242	4296.333	14490.613
Medicare OP Spending	1561.416	5323.631	2080.148	6085.013
Medicare Carrier Spending	3050.681	9175.846	3426.097	7087.399
Medicare HH Spending	6623.551	7488.791	5829.986	5957.003
Medicare SNF Spending	15446.217	14230.274	15047.545	13405.396
Medicare Other Spending	3550.166	4887.635	3892.715	5143.632

Table 2: Summary Statistics

	(1) All States	(2) Liberal	(3) Conservative	(4) MMC Loose	(5) MMC Strict
Effect on State Mean	0.837**	0.672**	0.590**	0.300**	0.714**
	(0.004)	(0.004)	(0.005)	(0.004)	(0.022)
Observations	1.87e+06	1.55e+06	1.23e+06	732660	260714
Treatment Observations	70,984	50,217	31,052	20,128	2,587

Standard errors in parentheses
$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Table 3: Regression Estimates of Effects of Moves on Spending

Difference in average log utilization				
	Median	25th	10th	5th
Overall	1.420	2.053	2.593	2.848
Patient	0.559	0.736	0.622	0.915
Place	0.861	1.317	1.971	1.934
Share of Difference				
Patients	0.394	0.358	0.240	0.321
Place	0.606	0.642	0.760	0.679

Table 4: Additive Decomposition

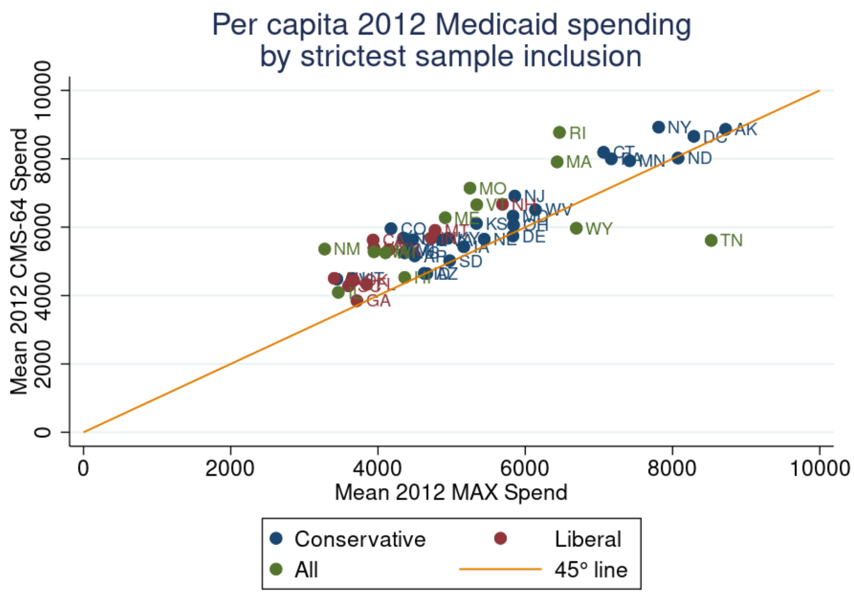
For Online Publication

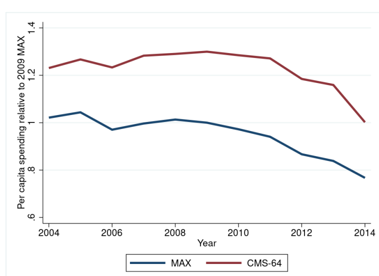
Appendix for:

Estimating Variation in Productivity Across State Medicaid Programs: Evidence from Dual-Eligibles

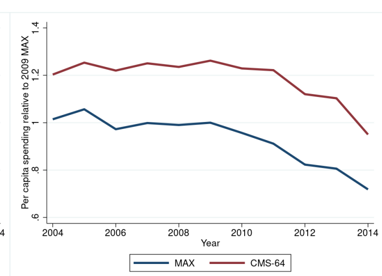
A Additional Figures

Appendix Figure A1: Cross-Sectional Comparison of Total Fiscal Spending in Medicaid MAX and CMS-64 Files in 2012

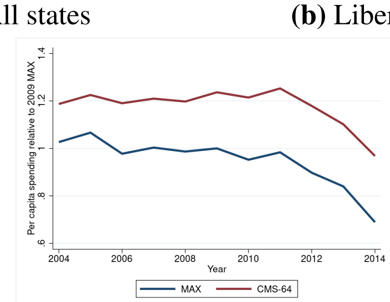




(a) All states



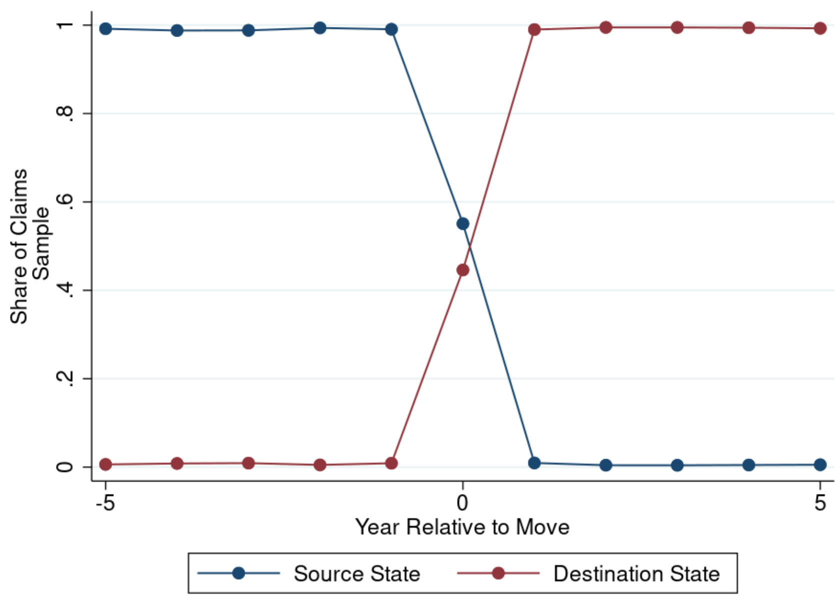
(b) Liberal states



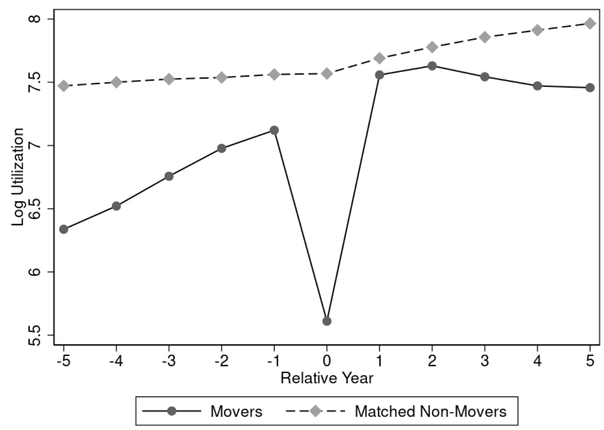
(c) Conservative states

Appendix Figure A2: Time Series Comparison of Total Fiscal Spending in Medicaid MAX and CMS-64 Files from 2004-2014

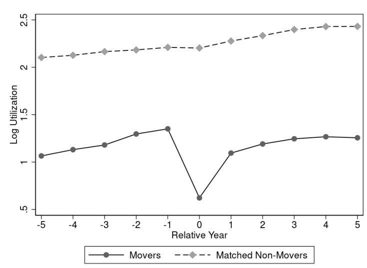
Appendix Figure A3: Share of Claims in Source and Destination States Pre- Vs. Post-Move



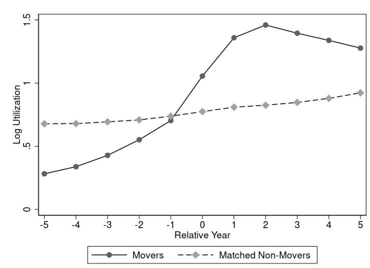
Appendix Figure A4: Log Spending over Relative Years among Movers and Non-Movers



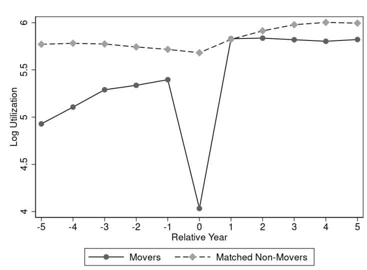
(a) Total spending



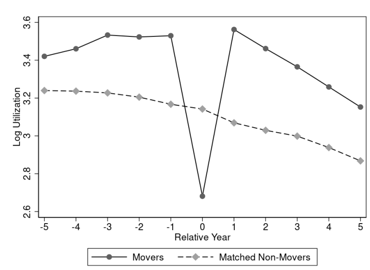
(b) LTSS spending



(c) LTC spending

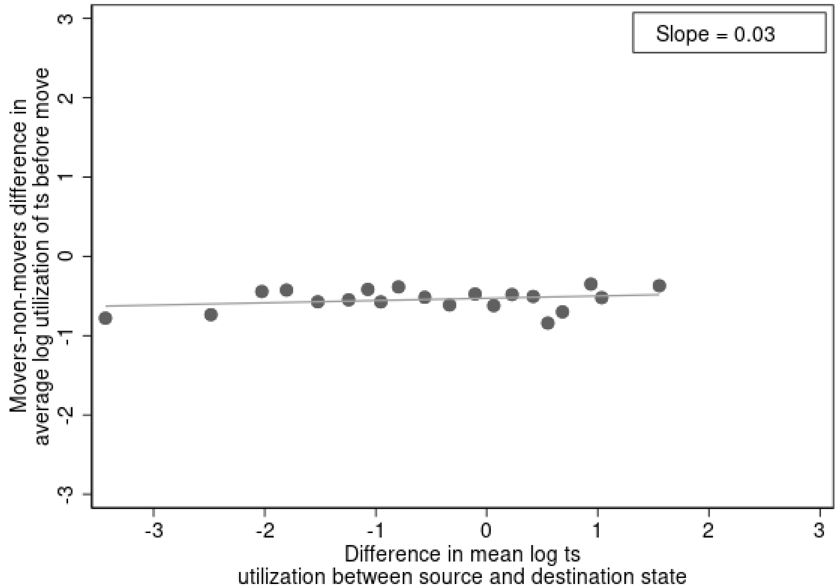


(d) Other primary payer

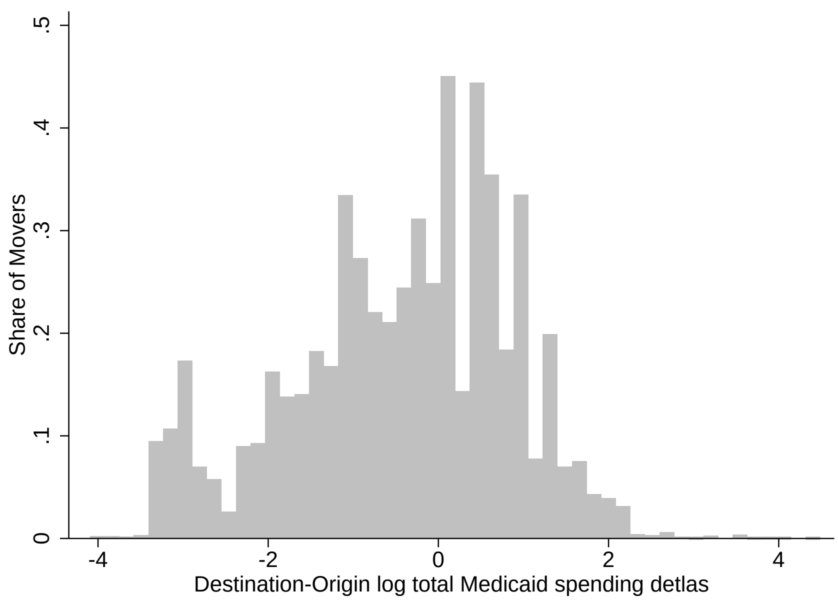


(e) Other secondary payer

Appendix Figure A5: Binscatter of Pre- Move Spending Versus Difference between Origin and Destination State Average Spending



Appendix Figure A6: Origin-Destination Differences in Total Spending Among Movers



B Additional Tables