Means-Tested Solar Subsidies*

Mark Colas
University of Oregon

Emmett Saulnier
University of Oregon

June 28, 2024

Abstract

We study the optimal design of income-contingent subsidies for residential solar panels. Using remotely sensed data on solar panel installations across the contiguous US and a border-discontinuity design, we estimate that the responsiveness of installation rates to subsidies is strongly decreasing in the local income level. Using these empirical elasticities, we estimate a model that embeds a solar panel installation decision into a dynamic consumption/savings framework with borrowing constraints. Counterfactual simulations reveal that switching to production-maximizing income-contingent subsidies leads to a three-fold increase in public funds received by low-income households and a 2.4% increase in national solar production. Means-tested subsidies are justified on both equity and efficiency grounds.

JEL Classification: H21, H23, Q42, Q48

Keywords: rooftop solar, subsidies, renewable energy

*We would like to thank David Evans, Stephie Fried, Bobby Harris, Joshua Graff Zivin, Kathleen Mullen, Ed Rubin, Chris Timmins. This work benefited from access to the University of Oregon high performance computing cluster, Talapas. This project was supported in part by an appointment to the Research Participation Program at the Water Economics Center, US Environmental Protection Agency, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the US Department of Energy and EPA. Any views expressed are solely those of the authors and do not necessarily represent the opinions of the US EPA.

1 Introduction

Subsidies for residential solar panels disproportionately benefit high-income households (Borenstein and Davis, 2016; Borenstein, 2017). Means-tested subsidies are an increasingly popular way to provide subsidies more equitably. As of 2023, 8 states have subsidy programs for residential solar which explicitly base eligibility or subsidy amounts on household income.[1] The number and size of these programs are expected to increase dramatically following the June 2023 launch of the Biden-Harris administration's Solar for All program, which pledged 7 billion dollars to fund low-income solar programs.[2] However, despite the increasing importance of income-contingent solar subsidies in US energy policy, to the best of our knowledge, no quantitative analysis exists on how to set these subsidies optimally.

This paper studies the equity and efficiency trade-offs associated with income-contingent subsidies for residential solar. We derive sufficient statistics for the cost-effectiveness of means-tested subsidies and estimate these sufficient statistics using border-discontinuity regressions. We use these empirical estimates to identify parameters in our structural model and then use our structural model to solve for optimal income-contingent subsidy schedules. We conclude that there are substantial equity and efficiency gains to instituting federal means-tested subsidies for rooftop solar panels.

To motivate our reduced-form analysis, consider a government that uses income-contingent subsidies to maximize solar production subject to a fiscal cost constraint. All else equal, the government will optimally provide subsidies to income groups with many *additional* households, that is, households who will be induced to install solar panels in response to a small subsidy increase. On the other hand, the fiscal costs of providing subsidies are increasing in the number of *non-additional* households: the households who already choose to install solar panels absent the subsidy increase. We show analytically that the ratio of the additional over non-additional households, as measured by the partial elasticity of solar production with respect to subsidies, can be used as a sufficient statistic for the cost-effectiveness of income-targeted subsidies. Specifically, we show that if this partial elasticity is decreasing (increasing) in income,

[1] https://www.solarreviews.com/blog/free-solar-panels-for-low-income-families. Several additional states also have community solar programs that target low-income households.

[2] Proponents of the program emphasized the program's equity benefits. Senator Bernie Sanders, the program's sponsor, said, "At a time when people are struggling to make ends meet, all while dealing with the existential threat of climate change, we must make residential rooftop solar a reality for low-income and working families that need it most. This $7 billion residential solar program... is a major step in the right direction."

then cost-neutral increases in the progressivity of the subsidy schedule will increase (decrease) total solar production.

We then empirically investigate how this partial elasticity varies across income levels. We use the Deepsolar database (Yu et al., 2018), which applies a novel machine-learning framework to satellite imagery across the contiguous US to measure the total residential solar panel area in each census tract, along with state-level subsidy data. Across a wide range of specifications, including border-discontinuity regressions and specifications that allow responsiveness to subsidies to vary nonparametrically in tract-level income, we consistently find that low-income tracts have partial elasticities greater than those of high-income tracts. Tracts with the median income level have a partial elasticity 15% to 43% higher than tracts at the 90th percentile. These results are robust to a battery of alternative specifications and suggest that means-tested subsidies could increase total solar production without increasing fiscal costs.

We then turn to a quantitative model of residential solar demand to evaluate counterfactual income-contingent subsidy schemes. Our framework embeds a homeowner's decision to install rooftop solar panels into a dynamic consumption/savings framework. Installing solar panels involves an upfront monetary cost but delivers subsidies and a stream of electricity production over the life of the panel. Households face borrowing constraints and, therefore, may not be able to fully smooth consumption if they choose to purchase solar panels. The model includes household heterogeneity in solar irradiance, preferences, and prices faced by households across space, as well as a rich quantification of the current federal and state subsidy schemes for solar panels, which accounts for differences in time profiles across which subsidies are paid and the nonrefundable nature of the Federal Investment Tax Credit.

We structurally estimate the model via indirect inference by using the remotely sensed data on residential solar installations from Deepsolar, as well as data on solar irradiance, electricity prices, subsidies, and income distributions across census tracts in the US. To achieve identification, we target our partial elasticity estimates from our border discontinuity regressions, as well as additional moments on solar panel installations across the income distribution and demographic groups. We show that our sparsely parameterized model matches both installation rates and elasticities of installations with respect to subsidies across the income distribution and is consistent with non-targeted quasi-experimental estimates of the responsiveness of solar installations with respect to prices and subsidies.

Our estimated structural model provides a rich framework that allows us to quantify the equity-efficiency consequences associated with various subsidy schedules. We first use the estimated model to analyze the effects of introducing federal income-contingent subsidies to the current subsidy scheme. Consistent with our reduced-form results, we find that introducing income-targeted subsidies for low-income households induces more electricity production per dollar of public funds than income-neutral subsidies. We show that the high number of non-additional households at high income levels is the primary driver of differences in cost-effectiveness across income groups: 50% of subsidies targeted for households at the 75th income percentile are received by households who would have installed panels absent these subsidies, compared to roughly 30% of households with incomes at the 25th percentile.

We then solve for the optimal income-contingent subsidies. We first consider a planner who chooses subsidies to maximize solar production without increasing fiscal costs. Despite the fact that this social objective places no weight on equity, we find that the optimal subsidies are highly progressive. The increase in progressivity associated with moving to these optimal subsidies and the resulting increase in installation rates for low-income households lead to a much more equitable distribution of public funds: the amount of solar subsidies received by households in the bottom income quartile triples, while the amount received by households in the top quartile is reduced by nearly half. The production-maximizing subsidy scheme also leads to a 2.4% increase in national solar production compared to income-neutral subsidies at no additional fiscal cost.

We then consider a planner who maximizes utilitarian welfare subject to a net cost constraint. The optimal welfare-maximizing schedule is even more progressive than the production-maximizing subsidies, as means-tested subsidies increase solar production while directing funds towards poorer households who have higher marginal utilities of income. This progressive subsidy scheme increases national solar production by 2.3%. Means-tested subsidies are justified on both equity and efficiency grounds.

We then conduct a series of robustness and sensitivity checks of our structural results. We examine how sensitive our findings are to 1) alternative assumptions on households' dynamic income process, 2) alternative levels of the borrowing limit, and 3) a government who maximizes environmental benefits rather than solar production. Across all specifications, we reach the same qualitative conclusions: optimal income-contingent subsidies are decreasing in income and switching to means-tested subsidies leads to substantial equity and efficiency gains. Finally, we perform a model-based decomposition to better understand the mechanisms which generate the positive relationship between income and installation rates. We find that borrowing constraints and the non-refundability of the Federal Investment Tax Credit the largest roles in generating the positive relationship between installations and income.

To the best of our knowledge, ours is the first paper to quantify the optimal income-contingent subsidies for solar panels. This focus on income-contingent subsidies differentiates our paper from other papers which use structural models to analyze the effectiveness of various types of income-neutral subsidies for solar panels (e.g., Burr (2014), De Groote and Verboven (2019), Snashall-Woodhams (2019), Langer and Lemoine (2022), Feger, Pavanini, and Radulescu (2022), Colas and Saulnier (2023)). We additionally contribute to this literature methodologically by 1) modeling households' dynamic consumption/savings decisions in an environment with borrowing constraints and 2) utilizing a utility function that exhibits nonzero income effects.[3] These features allow our model to replicate two patterns in the data that play a pivotal role in determining the returns to targeted subsidy increases: 1) installation rates across the income distribution and 2) elasticities of installations with respect to subsidies across the income distribution.

More broadly, we are also related to a literature which quantifies the distributional effects of energy policy (e.g., Bento et al. (2009), Borenstein (2012), Jacobsen (2013), Borenstein and Davis (2016), Fried, Novan, and Peterman (2018), Reguant (2019), Davis and Knittel (2019), Holland et al. (2019), Goulder et al. (2019), Morehouse (2021), Hahn and Metcalfe (2021), Linn (2022), Cahana et al. (2022), Fried, Novan, and Peterman (2022), and Dauwalter and Harris (2023)). We contribute to this literature by quantifying the dual equity-efficiency benefits in the case of residential solar subsidies. Finally, this paper is also related to several reduced-form papers estimating the responsiveness of solar installations to subsidies in the United States, which we discuss in Section 3.4. Related to these papers, we focus on how the responsiveness of installations varies as a function of household income.

[3] Feger, Pavanini, and Radulescu (2022) use detailed Swiss data to analyze the equity-efficiency trade-offs associated with solar panel cost subsidies and energy tariffs. They do not model savings and instead take a slightly more "reduced-form" approach where differences in installation behavior across income groups are generated by preference parameters that directly depend on household wealth. Snashall-Woodhams (2019) uses a dynamic model and data from California to solve for cost-minimizing subsidies that vary by electricity consumption type, solar energy potential, and location. Dorsey and Wolfson (2023) analyze differences in solar installation purchases across income and race groups and calculate differences in consumer surplus across demographic groups.

2 Reduced-Form Analysis

2.1 Cost-Effectiveness and Means-Tested Subsidies

We begin by deriving simple sufficient conditions for the cost-effectiveness of progressive subsidies in a general model. Individual households are associated with an income level $y \in \left[\underline{y}, \overline{y}\right]$. The government has access to a system of income-contingent subsidies for solar electricity production characterized by the function s, where $s(y)$ denotes the production subsidy available for households with income y.[4] Let $K_y\left(s(y)\right)$ be an increasing function which maps subsidies for income level y to total solar production from households of income level y. Further, let

$$Prod[s] = \int_{\underline{y}}^{\overline{y}} K_y\left(s(y)\right) dy$$

denote the functional which maps the subsidy function s to total electricity production and let

$$Cost[s] = \int_{\underline{y}}^{\overline{y}} K_y\left(s(y)\right) s(y) dy$$

denote the functional which maps s to fiscal cost.

Assume there are initially no income-contingent subsidies, as is the case federally, such that $s(y) = \bar{s}$ for all income levels y, where $\bar{s}$ is a nonnegative constant. We are interested in the implications of small changes to the subsidy function around this income-neutral baseline. Formally, let δs denote a *variation* to the function s such that subsidies received by any income level y change to $\bar{s} + \delta s\left(y\right)$, where $\delta s\left(y\right)$ represents an arbitrary infinitesimal change to subsidies.[5] We will focus on variations that are 1) cost-neutral and 2) progressive. Cost-neutral variations are those that lead to no change in fiscal cost, and progressive variations are those that are decreasing in income, such that subsidies become more generous for low-income households and less generous for high-income households.[6]

[4]For simplicity, we assume that the government only has access to these income-contingent production subsidies. In the structural model, we will include a rich model of subsidies for solar panels, including state and federal investment subsidies.

[5]Explicitly, we consider moving from the subsidy function $s\left(y\right) = \bar{s}$ to the alternative subsidy function $\tilde{s}(y) = \bar{s} + \epsilon g(y)$, where $g(y)$ is a function in y and where we take the limit as $\epsilon \to 0$.

[6]Formally, a cost-neutral variation is any variation δs such that

$$\int_{\underline{y}}^{\overline{y}} \frac{\delta Cost}{\delta s\left(y\right)} \delta s\left(y\right) dy = 0,$$

Proposition 1 provides a simple sufficient condition for when cost-neutral progressive subsidy variations lead to increases in solar production.

Proposition 1. *Define*

$$\eta(y) \equiv \frac{\frac{\partial K_y}{\partial s(y)}}{K_y} \tag{1}$$

as the "cost-effectiveness" of a subsidy increase for a given income level y*. If* η *is weakly decreasing (increasing) in income and* $\eta\left(\underline{y}\right) > \eta\left(\overline{y}\right)$ $\left(\eta\left(\underline{y}\right) < \eta\left(\overline{y}\right)\right)$*, then any cost-neutral progressive subsidy variation leads to a strict increase (decrease) in solar production.*

Proof. Appendix A.1 □

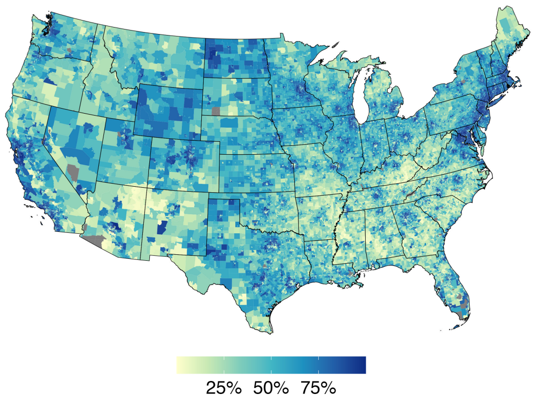
Proposition 1 shows that we can use measures of η across the income distribution as sufficient statistics for when moving to progressive subsidies can increase production: if η is decreasing in income, then cost-neutral progressive subsidies variations will lead to production *increases*. If η is decreasing in income, then these progressive subsidy variations lead to production *decreases*.

To see why this is the case, consider first the derivative $\frac{\partial K_y}{\partial s(y)}$ in the numerator of equation (1), which we refer to as the amount of "additional production." This value tells us how much solar production will increase if we increase subsidies by a small amount for a given income level. All else equal, it is more cost-effective for the government to increase subsidies for income levels where additional production is high, as subsidy increases for these households will lead to larger increases in total solar production. The denominator measures the amount of "non-additional production," the amount of production by panels installed absent that subsidy increase. If subsidies increase, households receive higher subsidies for this non-additional production, regardless of whether they increase their solar production. Thus, more non-additional solar production is associated with higher marginal costs to the government. Together, we can think of this ratio of additional over non-additional production as measuring the "bang for your buck" of a targeted subsidy increase since it measures the total change in solar production per dollar paid to non-additional households. Our goal in the coming sections is to empirically estimate how η varies across the income distribution.

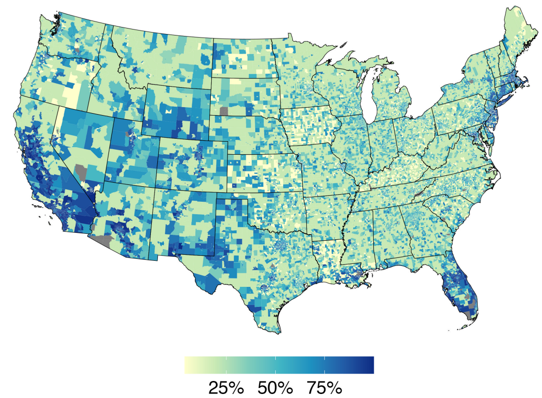
where $\frac{\delta Cost}{\delta s(y)}$ is the functional derivative of $Cost[s]$ with respect to $\delta s\,(y)$. A progressive variation is any variation δs such that

$$\left(y'' - y'\right)\left(\delta s(y'') - \delta s(y')\right) < 0$$

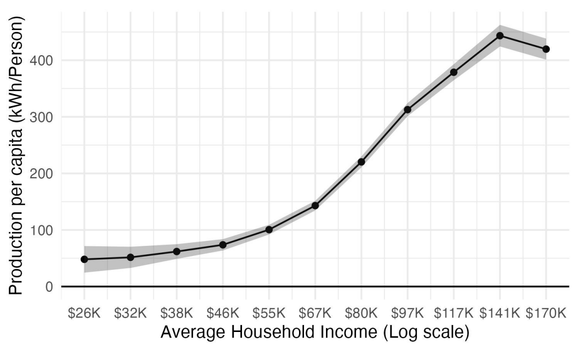
for $y'' \neq y'$.



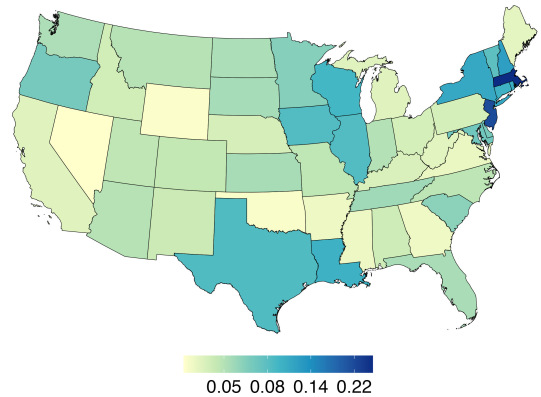
(a) Census tract income percentiles.



(b) Census tract residential solar production percentiles.



(c) Residential solar production per capita by income bin.



(d) State residential solar subsidies in \$/kWh.

Figure 1: Spatial variation in tract income, tract solar panel production, and state solar subsidies. Income comes from the 2015 5-year ACS, solar production from Deepsolar, and subsidies from Sexton et al. (2021). Panel (c) shows coefficients from regressing annual solar production per captia on 11 income bins evenly spaced in log income. Standard errors are clustered by state.

2.2 Data and Descriptive Results

We empirically estimate the cost-effectiveness of income-targeted subsidies using tract-level variation in solar production and state-level variation in subsidies. For this, we need a large sample of tract-level data on installations and income levels across the USA and state-level data on solar panel subsidies.

Solar panel installations We use the Deepsolar database (Yu et al., 2018) for tract-level residential solar panel installations and total panel area across the 48 states in the contiguous US. These data result from a deep-learning model trained to detect solar panels from satellite imagery captured in 2016, providing the first comprehensive and spatially fine measurement of solar panels.[7] In addition to providing the deep-learning model output, Deepsolar attaches several other variables for each census tract collected from many sources. These include solar irradiance from NASA Surface Meteorology and Solar Energy, state-average retail electricity prices from the EIA, and average income and other demographic data from the 2015 5-year American Community Survey (ACS).[8] We calculate solar production in each census tract by multiplying total residential panel area in a tract by its average solar radiation.

Panels (a) and (b) of Figure 1 show the spatial distribution of tract-level average income and total residential solar production. Solar production is concentrated in sunny areas such as the Southwest, Florida, and California and high-subsidy states in the Northeast. Meanwhile, income is highest along the East Coast, California coast, and surrounding major cities. Panel (c) depicts the relationship between tract-level average income and solar production per capita. Solar production is strongly increasing in income: the lowest income tracts, those with average income less than 38 thousand dollars, produce around 50 kWh of residential solar electricity per capita annually. The highest income bins, those with an average income over 117 thousand dollars, produce around 400 kWh of residential solar electricity per capita annually, eight times higher than the lowest income bins.

Solar subsidies We use measures of state-level solar panel subsidy generosity calculated by Sexton et al. (2021) using 2017 data from the Database of State Incentives for Renewables and Efficiency (DSIRE). They calculate these generosity measures as the total subsidies per kWh of production that an average-sized installation in each state is eligible for, accounting for federal and state investment tax credits, state production credits, property and sales tax rebates, Solar Renewable Energy Certificates, and other state-level subsidies.[9] We will refer to these measures as the "generosity" of subsidies in each state, which we use in the reduced-form estimation of how responsiveness to subsidies varies across income groups. We disaggregate into several different types of subsidies in the formal quantitative model that follows.

Panel (d) of Figure 1 shows the spatial distribution of state subsidies. Generally

[7] Alternatives rely on self-reported data (e.g., Open Solar Project) or do not cover the entire US (e.g., Tracking the Sun).

[8] Notably, we use population density, percent with a college degree, percent owner-occupied homes. Percent voting Democrat in the 2016 election comes from townhall.com.

[9] Unlike Sexton et al. (2021), we do not include net metering as a subsidy in our analysis as the government does not pay for it.

speaking, subsidies are most generous in the Northeast. Massachusetts has the highest subsidies in the country, at 28 cents per kWh. Meanwhile, several states that do not offer any subsidies in addition to the federal incentives have subsidies under 4 cents per kWh.

Border Discontinuities Our main empirical strategy is to use border-discontinuity regressions to estimate how the responsiveness of solar production varies across income levels. To motivate this strategy, we present descriptive evidence on how solar production levels in low- and high-income tracts change as we cross the border from a state with less generous subsidies to a state with more generous subsidies.

We define a tract's location relative to the nearest border as the positive distance to the border for tracts on the side of the border with more generous subsidies and the negative distance to the border for tracts on the side with less generous subsidies. We then categorize tracts into 10-mile-wide bins based on this location relative to the border, and regress log solar production per capita on state-border fixed effects, fixed effects for these location bins, and controls for population density.[10] We run these regressions separately for high- and low-income tracts, where tracts are categorized as "high-income" if their average income is in the top quartile of tract-level income and are labeled as "low-income" otherwise. We then plot the estimated location-bin fixed effects, which show conditional average production levels for low- and high-income tracts in narrow bandwidths around state borders.

Ideally, we would like to compare discontinuities in these average production levels at state borders to learn about how the responsiveness of solar production varies across income groups. One issue is that the distribution of high- and low-income tracts across state borders differ, and therefore, the number of observations associated with a given border will not be consistent between the regression with only high-income tracts and the regression with only low-income tracts.

This difference makes comparisons across high- and low-income tracts difficult because, as we show in Appendix B.1, high-income tracts tend to be located around borders with larger differences in subsidy generosity. We, therefore, reweight observations in each of the two regressions such that a given state border receives the same weight in both the regression with low-income tracts and the regression with high-income tracts.[11]

[10]This follows the approach used by Bayer, Ferreira, and McMillan (2007) to visualize how house prices respond to changes in school quality around school district borders.

[11]For example, New Hampshire and Massachusetts have very different subsidy generosity, and the

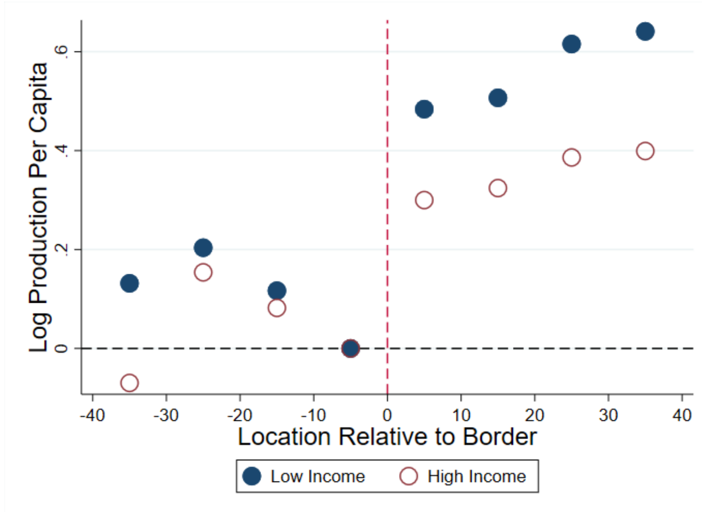


Figure 2: Border Discontinuities in Log Production per Capita. The graph plots estimated location-bin fixed effects from a regression log production per capita in high-income tracts (hollow circles) and low-income tracts (solid circles) on border fixed effects, location-bin fixed effects, and controls for population density. Positive values on the X-axis represent households on the side of the border with more generous subsidies, and negative values on the X-axis indicate the side of the border with less generous subsidies. Observations are reweighted such that the total sum of weights around each state border is equal to the total population of all tracts in that border region.

Figure 2 plots the estimated location-bin fixed effects for high-income and low-income tracts. Positive values on the X-axis represent tracts on the side of the border with more generous subsidies, and negative values on the X-axis represent tracts on the side with less generous subsidies.[12] We can see that log production rates for both groups increase sharply as we move to the side with more generous subsidies and that the increase in production is larger for low-income tracts: production rates increase by roughly 50% in low-income tracts compared to roughly 30% in high-income tracts.

border between New Hampshire and Massachusetts is surrounded mostly by high-income tracts. On the other hand, Arkansas and Mississippi have very similar subsidy levels, and mostly low-income tracts surround the border between Arkansas and Mississippi. Without this reweighting procedure, the discontinuity graph for high-income tracts would be more reflective of the response to the large subsidy discrepancy between New Hampshire and Massachusetts, while the discontinuity for low-income tracts would be more reflective of the small subsidy discrepancy between Arkansas and Mississippi. We reweight observations in each of the two regressions such that the sum of weights around each state border is equal to the total population of *all* tracts in that border region. Formally, let Pop_ℓ denote the population of a given tract ℓ, let $\overline{Pop}_\theta$ denote the total population in all tracts in the region around a given border θ, and let $\overline{Pop^I}_\theta$ denote the total population tracts of income group I around a given border θ. In our regression for tracts of income level I, we weight tract ℓ in border region θ by $\overline{Pop}_\theta \frac{Pop_\ell}{\overline{Pop^I}_\theta}$. We show the graphs without reweighting in Appendix B.1.

[12]The regressions omit the location-bin fixed effect for the location bin nearest to the border on the less generous subsidy side. Therefore, we can interpret these estimated location bin fixed effects as the conditional average of log production rates in a given location bin relative to this omitted bin.

These descriptive results suggest that production rates in low-income tracts may be more responsive to increases in subsidies than production rates in high-income tracts.

2.3 Empirical Strategy

We can rewrite our measure of cost effectiveness in equation (1) as a partial elasticity,

$$\eta(y) \equiv \frac{\frac{\partial K_y}{\partial s_y}}{K_y} = \frac{\partial \log K_y}{\partial s_y}. \tag{2}$$

If these partial elasticities are decreasing in income, then means-tested subsidies can increase solar production at no additional fiscal cost. Our reduced-form strategy is to estimate how these empirical partial elasticities vary across income levels using tract-level data on solar installations outlined above.

Concretely, we estimate various forms of the regression

$$\log K_\ell = \beta(Y_\ell) s_\ell + x_\ell' \boldsymbol{\gamma} + \varepsilon_\ell, \tag{3}$$

where K_ℓ is total solar production in census tract ℓ, Y_ℓ is average income in tract ℓ, s_ℓ is the generosity of subsidies available in tract ℓ, and x_ℓ is a vector of controls. The object of interest, $\beta(Y_\ell)$, gives the empirical partial elasticity of solar capacity with respect to subsidies for tracts with average income Y_ℓ. In practice, we will use several methods to parameterize how $\beta(\cdot)$ varies as a function of income.

Border Discontinuity Regressions Our first strategy is to assume cost-effectiveness is linear in log income and utilize a border-discontinuity approach.[13] Formally, let θ denote the nearest state border to a given tract and let Loc_ℓ denote the location of the tract relative to this border, again defined as the positive distance to the border for tracts on the side of the border with more generous subsidies and the negative distance to the border for tracts on the side with less generous subsidies.[14] Further, let the variable $\widehat{\log Y_\ell}$ denote "de-medianed" log income, calculated as log tract-level income

[13] Hughes and Podolefsky (2015) and Colas and Saulnier (2023) also use border-discontinuity approaches to estimate the effects of subsidies on solar panel installations. Neither paper estimates how the responsiveness of installations with respect to subsidies varies across income groups.

[14] Which of the two sides is normalized as the positive side does not affect any of the results as we always use border-specific polynomials. In Appendix B.1 we visualize the discontinuities in subsidy generosity and installation rates at state borders. In Appendix B.2 we show that household demographics do not exhibit discontinuities at state borders.

less the median log income level. Limiting our sample to tracts within 40 miles of state borders, we run regressions of the following form:

$$\log K_\ell = \beta_0^{\text{Disc}} s_\ell + \beta_1^{\text{Disc}} s_\ell \times \widehat{\log Y_\ell} + x_\ell' \boldsymbol{\gamma}^{\text{Disc}} + g_\theta\left(\text{Loc}_\ell\right) + \varepsilon_\ell^{\text{Disc}}. \tag{4}$$

The term $g_\theta\left(\text{Loc}_\ell\right)$ is a border-specific smooth flexible function in a tract's location relative to the border, and controls for unobservables with may affect solar production rates. In practice, we specify these functions as border-specific polynomials which vary from degree 0 polynomials, in which case $g_\theta(\cdot)$ is simply a border fixed-effect, to polynomials of degree 5.

The parameters of interest are β_0^{Disc}, which gives the empirical partial elasticity of solar production for tracts at the median income level, and β_1^{Disc}, which dictates how this partial elasticity varies with income. A value of $\beta_1^{\text{Disc}} < 0$ implies that the partial elasticity of solar production with respect to subsidies is decreasing in income, and therefore, means-tested subsidies can increase solar production without increasing fiscal cost.

Nonlinear Specifications Our second strategy is to use data on all tracts and estimate two specifications which allow for flexible non-linearities in this empirical partial elasticity.[15] The first of these specifications divides tracts into "bins" based on their average income level and estimates separate coefficients for each bin. Letting $\mathbb{1}\left(\ell \in \text{Bin}_b\right)$ denote that tract ℓ falls within income bin b, we estimate

$$\log K_\ell = \beta_0^{\text{Bin}} s_\ell + \sum_{b \neq b_0} \beta_b^{\text{Bin}} s_\ell \times \mathbb{1}\left(\ell \in \text{Bin}_b\right) + x_\ell' \boldsymbol{\gamma}^{\text{Bin}} + \varepsilon_\ell^{\text{Bin}}. \tag{5}$$

Thus, β_0^{Bin} gives the empirical partial elasticity associated with some base income bin (denoted by b_0), and the β_b^{Bin} coefficients tell us how partial elasticity in income bin b differs from that of the base income bin. In practice, we will set the base income bin as the bin corresponding to median tract-level income level such that β_0^{Bin} gives the partial elasticity at this median income level.[16]

[15] We do not have enough power to reliably estimate the parameters of these specifications using border discontinuity regressions. These discontinuity regressions rely on within-border variation in subsidies, and thus, identification of a non-linear effect of subsidies requires within-border variation in subsidies within narrow bins in income. Since we only have tract-level income available, there is insufficient variation in subsidies across all income bins.

[16] Estimating these coefficients relative to the median income bin also allows for easy statistical testing of whether cost-effectiveness differs relative to the median income bin.

Next, we estimate a model that allows for heterogeneity in cost-effectiveness with cubic B-splines,

$$\log K_\ell = \sum_{h=1}^{H} \beta_h^{\text{Spl}} s_\ell \times B_h(Y_\ell) + x_\ell' \boldsymbol{\gamma}^{\text{Spl}} + \varepsilon_\ell^{\text{Spl}}, \tag{6}$$

where $B_1(Y)$ to $B_H(Y)$ are standard basis functions for a cubic B-spline degree H. This specification balances allowing for arbitrary non-linearities while estimating relatively few parameters. The estimated partial elasticity for a given income level Y is a weighted average over the H spline coefficients, where the weights are given by the basis functions $B_h(Y)$.

Controls and Alternative Specifications All regressions include controls for tract-level income, electricity prices, electricity prices interacted with income, solar irradiance, population density and population density squared, and a battery of tract-level demographic measures. The border-discontinuity regressions include the border-specific polynomials introduced above while the nonlinear specifications include census region or census division fixed effects.

Note that we can only use tracts with positive solar production when estimating regressions with $\log K_\ell$ as the outcome. About a quarter of tracts have zero production in the Deepsolar data. In Appendix B.5, we use a strategy similar to that of Gillingham and Tsvetanov (2019) to account for these zero-production tracts, estimating a logit model for the effect of subsidies on the probability of a tract having positive installations. We call this an "extensive" margin effect, going from zero solar production in a tract to some positive amount of solar production.

2.4 Reduced-Form Results

Border-Discontinuity Regressions Table 1 reports parameter estimates from equation (4), the border discontinuity regression with cost-effectiveness linear in log income. Each column corresponds to a different specifications. Specifications vary the bandwidth around state borders (either 40 or 80 miles from state borders) and the degree of the polynomials in location relative to the state border (polynomials of degree 0, 3, and 5).[17]

[17] Appendix B.6 shows results for various other bandwidths and border-specific-polynomial degrees. The results are similar except for the smallest bandwidths.

Column (3) contains our preferred specification, with a 40-mile bandwidth and 3rd-degree border distance polynomials. The estimates imply that a one cent per kWh increase in subsidies is associated with a 4.3% increase in solar production for a tract at the median income level and that this elasticity is decreasing in tract-level income. The same one cent per kWh subsidy increase is associated with only a 3.4% increase in solar production per capita for a tract at the 90th percentile of the income distribution. In other words, the empirical partial elasticity of solar production with respect to subsidies for tracts at the median income level is 28.0 percent higher than the partial elasticity for tracts at the 90th percentile income level. The estimates in column (4), where we use an 80-mile bandwidth, are similar, implying that the partial elasticity at the median income level exceeds that at the 90th percentile by 27.7 percent.

The first two columns of Table 1 show results when we include polynomials of degree 0 in location relative to state borders, equivalent to including border fixed effects. These specifications result in elasticities that have a similar slope with respect to income but are slightly higher in levels than our main specification—a one cent per kWh increase in subsidies increases solar production by 6.0 to 6.5 percent for a tract at the median income. These partial elasticities for the median income tracts are 15.3 to 18.7 percent higher than those of tracts at the 90th income percentile.

Finally, columns (5) and (6) of Table 1 use fifth-degree border polynomials and again find similar results. In these specifications, the partial elasticity of solar production with respect to subsidies is 38.1 to 43.5 percent higher for tracts at the median income compared to those at the 90th percentile. Taken together, we find that the cost-effectiveness of solar subsidies is strongly decreasing in tract-level income.

One threat to identification in these border-discontinuity regressions is the potential for household preferences for solar panels to be discontinuous at state borders. This would be the case if, for example, households with a stronger preference for solar panels tended to locate on the side of the border with more generous subsidies. We investigate this type of sorting in Appendix B.2, where we look for discontinuities in household demographic characteristics around state borders. Our findings show no evidence of sorting around state borders.

Beyond subsidies, some states implement other programs designed to encourage solar installations, such as state-sponsored financing programs for solar installations or incentives for builders to incorporate solar panels into newly built structures. Another potential threat to identification is that these other state measures may lead to discontinuities in solar installation rates at state borders. In Appendix B.6, we address

Border Polynomial Deg.	0		3		5	
Bandwidth (mi)	40 mi	80 mi	40 mi	80 mi	40 mi	80 mi
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Subsidy	6.04***	6.53***	4.31**	4.74**	3.57*	4.02**
	(0.949)	(1.08)	(2.01)	(1.93)	(2.11)	(1.95)
Subsidy × Log Income	-1.50**	-1.92**	-1.76***	-1.92**	-2.02***	-2.07***
	(0.704)	(0.913)	(0.428)	(0.845)	(0.450)	(0.681)
Fit statistics						
Observations	20,187	30,410	20,187	30,410	20,187	30,410
R^2	0.48	0.49	0.54	0.55	0.19	0.56

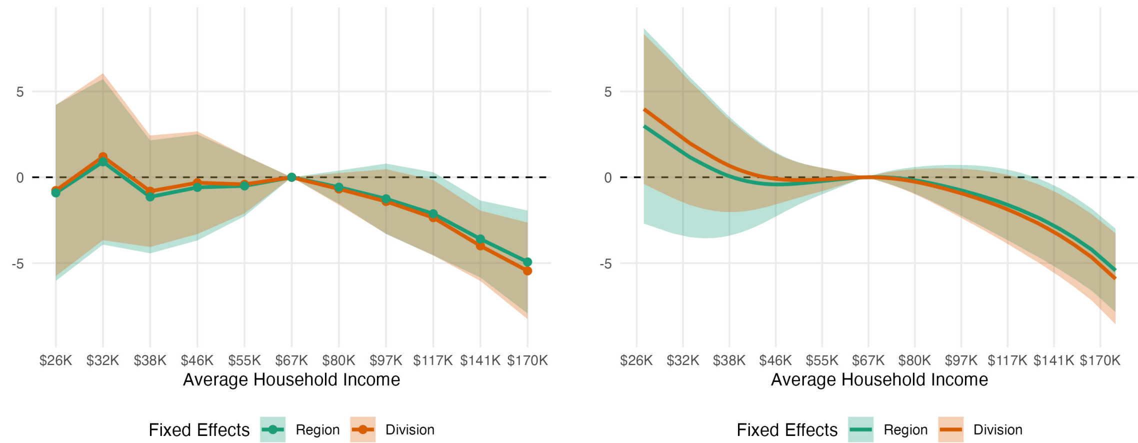
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 1: Effect of Subsidies on Log Production per Capita. Estimates of coefficients from Equation (4). Sample is limted to tracts within either 40 or 80 miles bandwidths to state borders. All regressions include controls for tract-level income, electricity prices, electricity prices interacted with income, solar irradiance, population density and population density squared, a battery of tract-level demographic measures, and border-specific polynomials in location relative to border. Clustered (State) standard-errors in parentheses

this by rerunning the border discontinuity regressions with additional controls for these other state-level programs aimed at increasing solar installations that are not included in our subsidy measures. Our results remain robust even with the inclusion of these additional controls.

Nonlinear Specifications Figure 3 demonstrates results from two approaches that allow the partial elasticity to vary non-linearly with income—using bins in log income and cubic B-splines over income levels—across specifications with either census region or division fixed effects.[18] Figure 3a shows estimates of the β_b^{Bin} from equation (5), where we divide tracts into 11 income "bins" based on their income level. The orange line shows estimates of β_b^{Bin} across income bins from a regression that uses census region fixed effects, and the green line shows estimates from a regression with census division fixed effects. Recall that β_b^{Bin} measures the partial elasticity in income bin b relative to the partial elasticity in tracts with the median income level. We obtain estimates of β_0^{Bin}, the parameter which gives the partial elasticity in tracts with the median income

[18]These models use data from all census tracts and include controls for tract-level income, electricity prices, electricity prices interacted with income, solar irradiance, population density and population density squared, and a battery of tract-level demographic measures.



(a) Cost-effectiveness using income bins.

(b) Cost-effectiveness using cubic B-splines over income.

Figure 3: Cost-effectiveness of solar subsidies across the income distribution. Panel (a) estimates a separate partial elasticity for 11 income bins, while panel (b) uses cubic B-splines. Both present estimates relative to a tract at the median income. We estimate cost-effectiveness to be 7.81 for the median income bin on the left and 7.36 at the median income using splines on the right when using census division fixed effects. The green line shows in each figure shows estimates from a regression which includes regions fixed effects and the orange line shows estimates from a regression which includes division fixed effects. All regressions include the controls variables listed in the text. Standard errors are clustered by state. Cubic B-splines have 7 knots evenly spaced based on population weighted income.

level of 6.0 in the regression with region fixed effects and 7.8 in the regression with division fixed effects.

In both specifications, the partial elasticity of production is relatively similar for tracts with income levels below the median income but drops steeply for tracts above the median income. Using the specification with region fixed effects, a tract at the 90th percentile of income (contained by the $117K income bin) has a partial elasticity of 3.9—implying that the partial elasticity for a tract at the median income is 54 percent higher than a tract at the 90th percentile. Similarly, the estimates from the regression with division fixed effects imply that the partial elasticity is 43 percent higher for a tract at the median income than a tract at the 90th percentile.

Figure 3b shows results where we use cubic B-splines to estimate cost effectiveness as a smooth and continuous function of income while still allowing for arbitrary nonlinearities. To maintain comparability to the bin specification, we normalize the estimated partial elasticities relative to that of the median income.[19] Again, the results from these specifications are similar to those of our other specifications. We estimate that a one-

[19] We use the variance-covariance matrix to calculate the standard errors of these relative estimates.

cent increase in subsidies leads to a 5.6 or 7.4 percent increase in solar production using region or division fixed effects, respectively. Meanwhile, a tract at the 90th percentile only increases solar production by 4.0 or 5.5 percent in response to the same one-cent subsidy increase. Thus, a median tract has a partial elasticity that is 33.5 to 39.2 percent higher than a tract at the 90th income percentile. Unlike when we use bins, the spline specification suggests that the partial elasticity may be considerably higher for the lowest-income tracts, though this increase is not statistically significant.

Regardless of the specification, we consistently find that the partial elasticity of residential solar production per capita decreases with income, with the steepest drop above median income. Our results suggest that cost-effectiveness may further increase in locations with below-median income relative to those with median income; however, this increase is not consistent across specifications and is not statistically significant.

Robustness In Appendix B.6, we examine the robustness of our reduced-form results to alternative outcomes (total production, total panels, panels per capita, installations, installations per capita), alternative controls, alternative bandwidths for border-discontinuity regressions, alternative means of estimating heterogeneity in the partial elasticity by income level, and a logit model that accounts for tracts with zero production. The results remain qualitatively the same in each case: the cost-effectiveness of residential solar subsidies is decreasing in income. Thus, a decrease in subsidies for high-income households and an increase in subsidies for low-income households could achieve increased residential solar production at the same fiscal cost.

3 Quantitative Model and Estimation

Our reduced-form results suggest that employing means-tested subsidies may allow policymakers to increase solar production without increasing fiscal costs. That analysis does not allow us to determine the magnitude of the benefits of introducing federal means-tested subsidies nor the optimal schedule of means-tested subsidies. Here, we construct and structurally estimate a partial equilibrium model of solar panel demand with borrowing constraints. Households in the model make a once-and-for-all decision whether to install solar panels, considering the lifetime costs and benefits. Households face borrowing constraints, meaning low-income households may not install panels despite the long-run benefits. We use the estimated model to evaluate the equity and efficiency consequences of introducing various income-contingent subsidy schemes and to quantify the optimal subsidy schedule.

3.1 Model

Households are indexed by i, and time is indexed by $t = 1, ...T$, which in our empirical setting will be years. In $t = 1$, households choose whether to install solar panels, considering the lifetime costs and benefits of installation.[20] Let $m_i \in \{0, 1\}$ indicate whether or not a household installs solar panels and let N_i be the number of solar panels household i installs conditional on installation.

Each household has access to solar panel technology that can produce a stream of solar energy of $\{A_{it}\}_{t=1}^{T}$ over time, where A_{it} represents the amount of electricity each panel installed by household i will produce in year t. This measure of solar production captures both differences in sunlight at their residence and depreciation of solar panels over time.

Budget Constraint Households who install panels pay an upfront installation cost but receive value from the electricity the panels produce over time. Let the function $p_j^{\text{Ins}}(N_i)$ denote the monetary cost of installing N_i panels and let p_j denote the price of electricity, where j indexes the state in which a household lives.[21] The total market value of electricity produced by household i in period t is therefore equal to $m_i N_i A_{it} p_j$. We can think of this as both the value of reducing the amount of electricity a household needs to purchase from the grid and the value of selling solar electricity back to the grid.

Households also receive subsidies for installing solar panels and for the electricity they produce. Let $s_i^{\text{Upfront}}(\cdot)$ denote the *upfront* subsidy household i would receive at the time of installation and let $s_i^{\text{Flow}}(\cdot)$ denote the *flow* subsidy the household i would receive each year over the life of the panel. When we take the model to the data, we will consider investment tax credits and sales tax rebates as upfront subsidies. We

[20] Burr (2014) use an optimal stopping model where households decide whether to install each period, exiting the market if they do choose to install. Gillingham and Tsvetanov (2019) presents evidence that households in most states do not treat installation as an install or wait decision.

[21] We assume that electricity can be purchased and sold back to the grid at this same price. For households in the 39 states with net metering, households receive the retail price of electricity for the electricity they sell back to the grid. In states without net metering, there may be different prices for electricity purchased by the household and electricity sold by the household. It would be straightforward to limit our estimation sample to states with net metering. Ito (2014) finds that consumers respond to average, rather than marginal, electricity prices.

allow upfront subsidies to depend on the cost of the installation, $p_j^{\text{Ins}}(N_i)$, to reflect that investment tax credits and sales tax rebates depend on the cost of installation. We also allow upfront subsidies to depend on a household's federal income tax burden to reflect that the federal investment tax credit is a nonrefundable tax credit, meaning that the amount a household receives cannot exceed the tax burden that they would otherwise owe.[22] Empirically, we categorize property tax rebates, renewable energy credits, and production tax credits as flow subsidies. We allow flow subsidies to depend on installation costs to reflect that property tax rebates often depend on the cost of installation and on A_{it} to reflect that renewable energy credits and production tax credits depend on the amount of electricity produced.

The household's budget constraint in year $t = 1$ is given by

$$c_{i1} + a_{i2} + m_i p_j^{\text{Ins}}(N_i) = y_i - \tau(y_i) + (1+r)\, a_{i1} + \\ m_i \left(N_i A_{it} p_j + s_i^{\text{Upfront}} \left(p_j^{\text{Ins}}(N_i), \tau(y_i) \right) + s_i^{\text{Flow}} \left(p_j^{\text{Ins}}(N_i), A_{it} \right) \right), \quad (7)$$

where c_{i1} is consumption of the numeraire good in period $t = 1$, a_{i2} is the amount the household saves for the following period, y_i is household income, $\tau(y_i)$ is federal income tax burden (not including federal solar credits), a_{i1} is initial assets, and r is the real interest rate. Due to data limitations, we assume that income for each household is constant over time. We consider a version of the model with a stochastic income process in Section 5.1.

After the first period, households continue to make consumption-savings decisions and receive electricity and flow subsidies from installed solar panels. Additionally, while federal tax credits are nonrefundable, excess credits can be carried over to the following years. We can write the household's budget constraint for $t > 1$ as

$$c_{it} + a_{it+1} = y_i - \tau(y_i) + (1+r)\, a_{it} + \\ m_i \left(N_i A_{it} p_j + s_i^{\text{Flow}} \left(p_j^{\text{Ins}}(N_i), A_{it} \right) + s_{it}^{\text{Carry}} \right), \quad (8)$$

where s_{it}^{Carry} gives the value of any federal tax credits that have been carried over from previous years.

Households face a borrowing constraint in each period. We write this as

[22] Kiribrahim-Sarikaya and Qiu (2023) uses data from Phoenix, Arizona to analyze the effects of the nonrefundable nature of the federal income tax credits on solar installation rates across the income distribution.

$$a_{it+1} \geq \bar{a}, \tag{9}$$

where $\bar{a}$ is the exogenously-given level of minimum assets a household must maintain. We set $\bar{a} = 0$ in our baseline specification. We consider alternative borrowing limits in Section 5.2. Our results are not sensitive to the value we set for $\bar{a}$.

It is important to highlight that we abstract away from changes in subsidies and prices over time. We can therefore view our model as capturing the forward-looking installation behavior for households who expected subsidies and prices to remain at their current levels.[23] In this regard, our model differs from those presented in De Groote and Verboven (2019) and Langer and Lemoine (2022) in that is not well suited to analyze how dynamic subsidy paths will affect the timing of installations. Instead, the main goal of our model is capturing differences in installation behavior across the income distribution for households facing a given set of subsidies.

Utility Households' lifetime utility is given by

$$\sum_{t=1}^{T} \beta^{t-1} \frac{(c_{it})^{1-\gamma}}{1-\gamma} + m_i \phi_i,$$

where $\beta = \frac{1}{1+r}$ is the household discount rate, γ is a preference parameter, and ϕ_i gives the nonpecuniary benefit of a solar installation for household i, reflecting inconvenience costs or other individual preferences for installing solar panels. We specify ϕ_i as

$$\phi_i = \phi_0 + \phi_{\text{Coll}} X_\ell^{\text{Coll}} + \phi_{\text{Pol}} X_\ell^{\text{Pol}} + \sigma \epsilon_i,$$

where X_ℓ^{Coll} is the fraction of individuals with a college education in the census tract in which the household lives, X_ℓ^{Pol} is the fraction of voters in the county who voted Democrat in the 2016 presidential election and ϵ_i is a logit preference draw with scaling parameter σ. Let $\bar{\phi}_i = \phi_i - \sigma\epsilon_i$ denote the portion of non-pecuniary utility that does not contain an idiosyncratic component.

[23] Hughes and Podolefsky (2015) and Anderson, Kellogg, and Sallee (2013) find that consumers do not correctly forecast the extent to which prices change over time and expect future prices to be similar to current prices.

Installation Probabilities Previous research has found that the number of panels per installation does not strongly correlate with rooftop solar subsidies.[24] We therefore abstract away from the intensive margin decision and parameterize N_i as a reduced-form function of household income and tract-level characteristics. We estimate the parameters of this reduced-form function jointly with the other structural parameters using indirect inference. We provide additional details in Appendix C.2.

Each household makes a discrete choice over whether to install solar panels and then makes consumption-savings decisions. Given that ϵ_i has a logit distribution, the probability household i installs solar panels is given by

$$P_i = \frac{\exp\left(\frac{1}{\sigma}\sum_{t=1}^{T}\beta^{t-1}\frac{\left(c_{it}^{m=1}\right)^{1-\gamma}}{1-\gamma} + \bar{\phi}_i\right)}{\exp\left(\frac{1}{\sigma}\sum_{t=1}^{T}\beta^{t-1}\frac{\left(c_{it}^{m=1}\right)^{1-\gamma}}{1-\gamma} + \bar{\phi}_i\right) + \exp\left(\frac{1}{\sigma}\sum_{t=1}^{T}\beta^{t-1}\frac{\left(c_{it}^{m=0}\right)^{1-\gamma}}{1-\gamma}\right)},$$

where $c_{it}^{m=1}$ and $c_{it}^{m=0}$ give the household's optimal consumption level in period t conditional on installing and not installing solar panels, respectively.

The parameter γ plays a crucial role in our analysis as it dictates how the elasticity of installations varies across income levels. To see this, note that partial elasticity of installation with respect to upfront subsidies s^{Upfront} is given by

$$\frac{\partial \log P_i}{\partial s^{\text{Upfront}}} = \frac{\left(c_{i1}^{m=1}\right)^{-\gamma}}{\sigma}\left(1 - P_i\right), \tag{10}$$

where $c_{i1}^{m=1}$ gives the household's optimal consumption choice in period 1 conditional on installing panels. We provide a derivation of equation (10) in Appendix C.1. Households with high income and asset levels will generally have higher values of $c_{i1}^{m=1}$. If γ is large, households with high income and asset levels will be less responsive in their installation decisions, all else equal.[25] This lower responsiveness is because a larger value for γ means the marginal utility of consumption decreases more rapidly as consumption levels increase.

Additionally, a higher installation probability P_i decreases responsiveness to subsidies. This decrease reflects that households with a higher P_i are more likely to be

[24] Colas and Saulnier (2023) find that installation size is not sensitive to monetary subsidies, but the probability of installations is highly responsive to monetary subsidies. They conclude that accounting for the extensive margin installation decision is much more important than the intensive margin decision of the number of panels to install.

[25] To see this, note that if $\gamma = 0$, the partial elasticity of installation across households will not depend on $c_{i1}^{m=1}$, and variation across households will only be due to differences in installation probability P_i.

inframarginal—their installation decision is unaffected by a marginal subsidy change. Finally, the parameter σ determines the overall level of the partial elasticity of installations with respect to subsidies. A larger value of σ, representing stronger idiosyncratic preferences for solar installations across all households, implies that households will be less responsive to subsidies in their installation decision.

3.2 Data

For structural estimation, we combine the data on subsidies, electricity prices, and residential solar installations described in Section 3.2 with income and homeownership data from the ACS and solar irradiance data from Google Project Sunroof. Here, we give an overview of the main data sources that we do not use in our reduced-form analysis. We provide additional details on the data we use for structural estimation in Appendix C.2.

Solar Potential We use data from Google Project Sunroof (GPS) to construct the solar potential for panels installed by each household. GPS applies a machine-learning framework to satellite imagery and provides measures of solar production capacity per panel at the tract level, accounting for local weather conditions, rooftop sizes, and shading. We assume a household's yearly solar potential for newly installed panels, A_{i1}, is equal to the mean household solar potential in the GPS data for the household's tract. We assume solar panel efficacy depreciates by a constant rate of 0.5% each year before fully depreciating after 20 years.[26]

Installation Prices We assume installation prices are given by the function $p_j^{\text{Ins}}(N_i) = p_j^{0,\text{Ins}} + N_i p_j^{1,\text{Ins}}$, where $p_j^{0,\text{Ins}}$ is a fixed cost and $p_j^{1,\text{Ins}}$ is a per-panel cost. We take estimates of $p_j^{0,\text{Ins}}$ and $p_j^{1,\text{Ins}}$ from Colas and Saulnier (2023), who estimate these pricing functions using data from Tracking the Sun, a project collecting data on solar panel installations. As Tracking the Sun does not cover all states within the US, they assume that pricing functions are common within each Census region.

Income and Initial Assets Simulating our model requires the household income distribution for homeowners across the United States. For this, we construct tract-level income distributions for homeowners by combining 1) tract-level data on average

[26] Jordan and Kurtz (2013) find a median degradation rate of 0.5% in their review of the literature on depreciation rates of solar panels.

household income, Gini coefficient, and number of households from the ACS, and 2) household-level data on homeownership and income from the ACS. We describe this procedure in detail in Appendix C.3.

An additional empirical issue is that we do not observe initial assets, a_{i1}. We treat initial assets as a latent variable and estimate a probability distribution over this latent variable for each household simultaneously with the rest of the model's structural parameters. Further details of this procedure are in Appendix C.4. We then integrate over the distribution of the latent variable when simulating model outcomes.

3.3 Estimation

Our primary strategy is to estimate the model by indirect inference, where we target regression coefficients from our reduced-form results and moments describing the distribution of installations across income and demographic groups. We first compute a set of "auxiliary models" that describe installation behavior in the data. Then, given a vector of structural parameters, we simulate the structural model and calculate the auxiliary models with simulated data. We repeat this procedure for different values of structural parameters and search for the parameters such that the auxiliary models computed from the model match those from the data.

Formally, let $\bar{\beta}$ denote the vector of auxiliary model parameters we estimate in the data and let $\hat{\beta}(\Theta)$ denote the same auxiliary model parameters computed from the structural model given an arbitrary vector of structural parameters denoted by Θ. The estimated vector of structural parameters is given by

$$\hat{\Theta} = \text{argmin}_{\Theta} \left\{ \left(\hat{\beta}(\Theta) - \bar{\beta} \right)' W \left(\hat{\beta}(\Theta) - \bar{\beta} \right) \right\},$$

where W is a weighting matrix.

We use the following auxiliary models[27]:

AM 1: Border discontinuity regression coefficients from estimation of equation (4). Specifically, we regress tract-level log production on subsidies and subsidies interacted with log average income for tracts within 40 miles of a border with 3rd degree polynomials in distance from each border, controlling for the control variables from column 3 of Table 1.

[27] We show the full set of auxiliary model parameters we use in Table A8 in Appendix C.6.

AM 2: Coefficients from regressions of tract-level installations per capita on tract-level average income.

AM 3: Average panels per capita by income quintile.

AM 4: Coefficients from regressions of tract-level installations per capita on tract-level demographics.

AM 5: Coefficients from a regression of tract-level average number of panels per installation on tract-level characteristics.

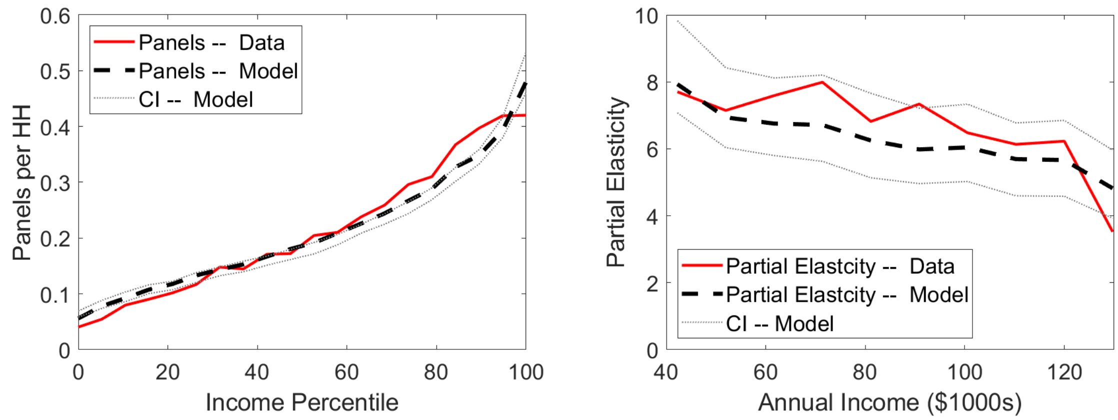
The parameters are well-identified. As previously noted, the parameter σ dictates the overall partial elasticity of installations with respect to subsidies across all households, while γ dictates how these elasticities vary with household income and assets. "AM 1" and "AM 2" jointly identify these two parameters. "AM 1" describes how the elasticity of installations with respect to subsidies varies by income, and "AM 2" describes how installations vary with income. The parameters ϕ_0, ϕ_{Coll}, and ϕ_{Pol}, which determine the nonpecuniary benefits of installations, are identified by "AM 4", which describes how installations vary with tract characteristics. Finally, "AM 5" pins down the parameters that determine how the size of installations varies by demographic group.

3.4 Model Fit

We present the full set of estimated parameters in Appendix C.5. The estimated model fits both targeted and untargeted moments well. We give an overview of model fit here and leave additional results in Appendix C.6.

Targeted Moments Figure 4a plots panels per household across percentiles of tract-level income in the model and data. The red line shows the panels per household in the data, while the black dashed line shows the simulated panels from the estimated model. The two dotted lines give the 95% confidence interval of the simulated moments, which we calculate by re-estimating the model and re-simulating outcomes using 100 bootstrap samples. The model does an excellent job of matching differences in the number of solar panels per household across the entire income distribution.

As discussed in Section 2, variation in the partial elasticity of solar production across the income distribution is critical for determining the cost-effectiveness of subsidies by



(a) Panels: Income

(b) Partial Elasticity

Figure 4: Panel (a) shows panels per household across percentiles of tract-level income in the model and data. The red line shows the panels per household in the data, while the black dashed line shows the simulated panels from the estimated model. The two dotted lines give the 95% confidence interval of the simulation moments. Panel (b) shows partial elasticities across income levels in the model and data. These are estimated using the same set of controls as Figure 3 with census division fixed effects. The two dotted lines give the 95% confidence interval of the simulated moments, which we calculate by re-estimating the model and re-simulating outcomes using 100 bootstrap samples.

income level. To assess the model fit in this dimension, we estimate the relationship between log tract-level installations and subsidy generosity across income groups in the model and the data. We again divide tracts into income bins based on the average income level in each census tract. Let $\mathbb{1}(\ell \in \text{Bin}_b)$ denote that tract j falls within income bin b. Using both actual data and simulated data from the model, we run regressions of the form

$$\log K_\ell = \sum_{b=1} \beta_b^{\text{Fit}} s_\ell \times \mathbb{1}(\ell \in \text{Bin}_b) + x_\ell' \gamma^{\text{Fit}} + \varepsilon_\ell^{\text{Fit}}, \tag{11}$$

where K_ℓ is the total solar production in tract j, s_ℓ denotes the subsidy generosity measure from Sexton et al. (2021) which we used in our reduced-form analysis, and b indexes income groups. We use the same controls as in Figure 3, with census division fixed effects. The β_b^{Fit} parameters, therefore, measure the partial elasticity of solar production with respect to subsidies for households income bin b. Figure 4b plots the estimates of these parameters for the model and the data. The model fits the empirical partial elasticity of installations with respect to subsidies across income levels well.

Untargeted Moments We now compare simulated results from the model to data not targeted in estimation to further assess model validity. The 2015 Residential Energy Consumption Survey (RECS) has income and solar installation status data for 5,700

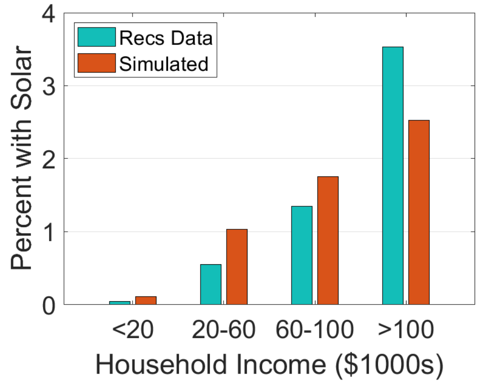


Figure 5: Untargeted Moments: Comparison to RECS installation rates. The bars show the percent of households in each income bin who have solar panels in the RECS microdata and the model simulation. RECS has income data in 8 income categories. We combine categories such that the bins in the figure roughly correspond to income quartiles in 2015.

households across the United States, along with data on other energy-related goods and behaviors. Figure 5 shows the percent of households with solar panels by income quartile from our model simulations and the RECS data.[28] The fit is reasonably good, though our model does predict lower installation rates for high-income levels than in the RECS data.

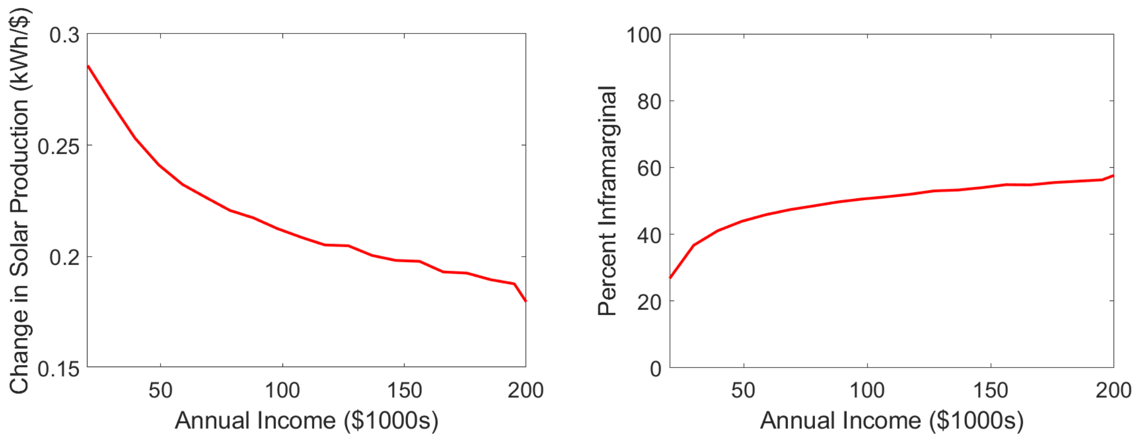
Comparison to Existing Literature Several other studies use diverse methods and empirical applications to credibly estimate the effect of subsidies on household demand for solar panels (Hughes and Podolefsky, 2015; Crago and Chernyakhovskiy, 2017; Gillingham and Tsvetanov, 2019; Colas and Saulnier, 2023). We use our estimated model to reproduce results from those studies, the details of which are in Appendix C.7. For each of these four studies, our replication is consistent with the results of the respective study.

4 Counterfactuals

4.1 Introducing Income-Contingent Subsidies

Before solving for the cost-minimizing subsidies, we first analyze the cost-effectiveness of introducing small income-contingent subsidies to the current subsidy scheme. Specifically, we calculate the additional solar capacity per dollar of public funds associated

[28]Income in RECS data is presented in income categories. We combine categories such that the bins in the figure roughly correspond to income quartiles in 2015.



(a) Additional production per dollar

(b) Percent of installations non-additional

Figure 6: Cost-effectiveness of introducing income-contingent subsidies to the current subsidy scheme. Panel (a) shows the change in solar capacity per dollar of additional fiscal cost associated with introducing income-contingent subsidies to each income group. Panel (b) shows the percent of income-contingent subsidy funds received by non-additional households.

with income-contingent subsidies. To calculate this, we divide households into 20 income groups. We then simulate installations 1) given the current system of subsidies and 2) where we also offer small, targeted subsidies for households of a given income group. We calculate the additional solar capacity per dollar for that income group as the increase in solar capacity divided by the increase in fiscal cost. We repeat the process for all income groups. We assume these income-contingent subsidies are upfront subsidies: they are paid in full at the time of installation. We compare the efficacy of upfront subsidies to flow subsidies in Section 4.4.

The results are presented in Panel (a) of Figure 6. Solar capacity per dollar of public funds is decreasing in household income. Introducing subsidies targeted at households with an income of $40,000 leads to an increase of 0.25 kWh of solar electricity per additional dollar of subsidies. On the other hand, subsidies targeted at high-income households are barely 70% as cost-effective: subsidies targeted at households with income over $200,000 lead to 0.18 kWh of solar electricity per dollar.

As highlighted in Section 2, a key determinant of the cost-effectiveness of introducing income-contingent is the number of non-additional households relative to additional households. To illustrate how this relationship varies across the income distribution, we calculate the percent of targeted subsidies for each income group received by households who choose to install panels absent the subsidy increase. Panel (b) of Figure 6 shows the results. The percentage of non-additional households is strongly increasing in income. For households with an income of $40,000, roughly 40% of targeted subsidy funds go to households who would already install solar panels absent the targeted subsidies. The percent of non-additional households is over twice this for high-income households: over 55% subsidies for households with income over $200,000 are received by households who would already install panels absent the subsidy increase.

4.2 Production-Maximizing Subsidies

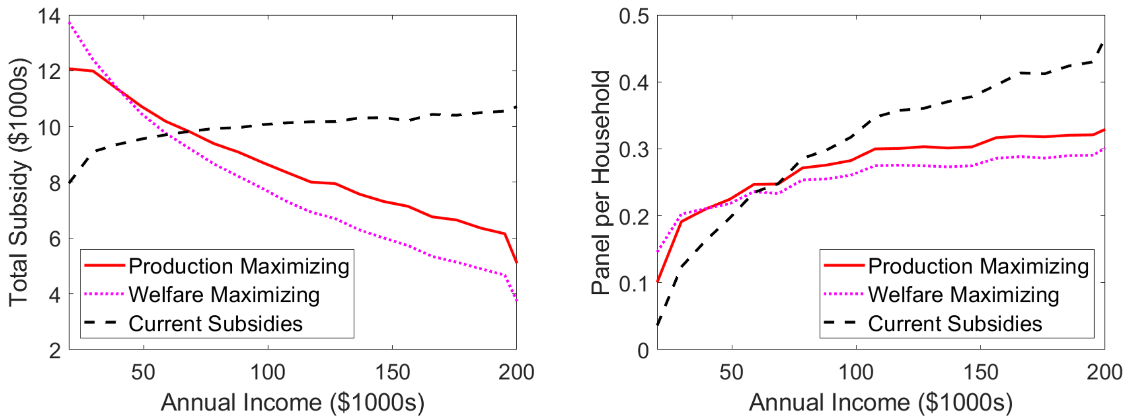
We consider a federal government that chooses a national income-contingent subsidy schedule to maximize total solar production subject to the constraint that total government spending must be no greater than spending under current subsidies. We hold all other state and federal subsidies constant and again assume that the government pays income-contingent subsidies upfront.[29] Note that since this objective does not account for household utility, we have removed the equity rationale for means-tested subsidies as there is no motive for the government to redistribute from rich to poor households. In Appendix C.10, we formalize the government's problem and derive its first-order conditions. To solve for the optimal subsidies, we numerically calculate the system of income-contingent subsidies that satisfy these first-order conditions.

The results are displayed in Table 2 and in Figure 7. Figure 7a presents the production-maximizing subsidies (solid red line) and the current subsidies (black dotted line) as a function of income. Specifically, each line shows the average present discounted value of subsidies a household of a given income level would receive conditional on installing solar panels.[30] Moving to the production-maximizing subsidies involves increasing subsidies for low-income households and decreasing subsidies for high-income households. Under the production-maximizing schedule, households in the first income quartile receive $12,000 in subsidies for an installation. In contrast, households in the top income quartile receive less than $7,000.[31] Figure 7b and Panel II of Table 2 show simulated installations across the income distribution under both subsidy schemes. Installations increase by nearly 50% for households in the first income quartile

[29]We compare the efficacy of upfront subsidies to flow subsidies in Section 4.4. We do not evaluate the optimality of the overall level of subsidies. Colas and Saulnier (2023) find that subsidy levels are suboptimally high in nearly all US states.

[30]Current subsidies are increasing in income for two main reasons. First, because the Federal Investment Tax Credit is nonrefundable, households with low income tax burdens cannot receive the full value of an upfront subsidy. Second, many subsidies pay a fraction of the cost of installation, and higher-income households tend to install larger, and therefore more expensive, solar panel systems.

[31]Echoing the results from the previous section, subsidies are generally decreasing in income except for the very lowest income levels. This is because many of the lowest-income households are likely to be severely borrowing constrained and unable to install solar panels even when subsidies increase.



(a) Average Subsidy

(b) Panels per Household

Figure 7: The figure on the left shows the average present discounted value of subsidies received for an installation across income levels under various subsidy schemes. The figure on the right shows the average panels per household across income levels under these same subsidy schemes.

while installations of households in the top income quartile decrease by roughly 35%.

As shown in Panel III of Table 2, these changes in subsidies and the profile of installations lead to a much more equitable distribution of public funds. Switching from the current subsidy schedule to the production-maximizing subsidy schedule triples the amount of solar subsidies received by households in the bottom income quartile, from 5.0% of total subsidies to 16.0%. On the other hand, funds received by households in the top income quartile drop by nearly half, from 47.2% to 25.9% of total subsidy payments.

Panel IV presents the relative total solar production of the production-maximizing subsidies. The production-maximizing subsidies increase total solar production by 2.4% relative to current subsidies with no increase in fiscal cost.

4.3 Welfare-Maximizing Subsidies

We now solve for the schedule of income-contingent subsidies that maximizes the sum of all households' lifetime utility subject to the constraint that net costs, which we define as total fiscal costs less environmental benefits, must not exceed the current level. To calculate environmental benefits, we estimate NERC-region-level marginal damages of electricity production during daytime hours using the estimation strategy and data from Holland et al. (2020). Our estimates measure the environmental damages of greenhouse gases and pollutants that fossil-fuel power plans would otherwise emit. We provide additional details in Appendix C.8. We formalize the maximization problem and present the first-order conditions in Appendix C.12.

	(1) Baseline	(2) Prod Max	(3) Utility Max
I. Production per HH			
Income Q1	13.2	33.4	46.3
Income Q2	48.5	64.3	65.2
Income Q3	80.2	80.4	76.0
Income Q4	123.5	96.7	88.4
Overall	68.2	69.8	69.7
II. Subsidy Generosity ($1000s)			
Income Q1	8.1	12.0	13.6
Income Q2	9.3	11.4	11.5
Income Q3	9.8	9.8	9.3
Income Q4	10.4	6.9	5.7
III. % of Public Funds Received			
Income Q1	5.0%	16.0%	24.2%
Income Q2	17.8%	27.9%	28.6%
Income Q3	30.0%	30.1%	27.1%
Income Q4	47.2%	25.9%	20.2%
IV. Relative Production	100.0	102.4	102.3

Table 2: Panel I shows the average annual solar capacity in kWh per household in each income quartile. Panel II shows the average present discounted value of subsidies a household from each income quartile would receive for a solar installation. Panel III shows the percentage of solar subsidies received by each income quartile. We measure these subsidies as the present value of all state and federal subsidies received by households in a given income quartile as a fraction of the total amount received across all households. Panel IV shows total solar production. We scale total production under the baseline simulation to 100.

The results are displayed in Column 4 of Table 2 and in the magenta dotted lines of Figure 7. The utility-maximizing subsidies are strongly decreasing in income, as the decreasing marginal utility of income gives the government an incentive to redistribute resources from higher-income households to lower-income households. Households in the first quartile of the income distribution receive 13.6 thousand dollars on average for a solar installation, while households in the top income quartile receive less than 6 thousand dollars. Switching to this welfare-maximizing subsidy scheme leads to a 2.3% increase in solar production.

4.4 Upfront Vs. Flow Subsidies

We now compare the efficacy of upfront and flow subsidies at inducing installations across the income distribution. We first simulate an income-neutral increase in upfront subsidies, denoted by $\Delta s^{\text{Upfront}}$. We then simulate an income-neutral increase in flow subsidies denoted by Δs^{Flow}. For the sake of comparison, we choose both $\Delta s^{\text{Upfront}}$ and Δs^{Flow} such that the change in fiscal cost associated with each subsidy increase is equal to 10% increase in total fiscal costs.

The results are shown in Figure 8. The solid red line shows the percent change in installations across the income distribution associated with increasing upfront subsidies by $\Delta s^{\text{Upfront}}$. Consistent with our other results, we find that the upfront subsidy increase leads to a much larger increase in installations for lower-income households. Installations for households in the first income decile, for example, increase by 29%, while installations for households in the top income decile increase by 3%.

The black dotted line shows the effect of increasing flow subsidies by Δs^{Flow}. The increase in installations is significantly muted at the bottom of the income distribution: installations for households in the bottom decile increase by only 8%. This occurs because, unlike upfront subsidies, flow subsidies do little to alleviate short-term liquidity constraints for low-income households. Altogether, these results suggest that upfront subsidies are both more cost-effective overall and more effective at increasing installations for low-income households than flow subsidies.

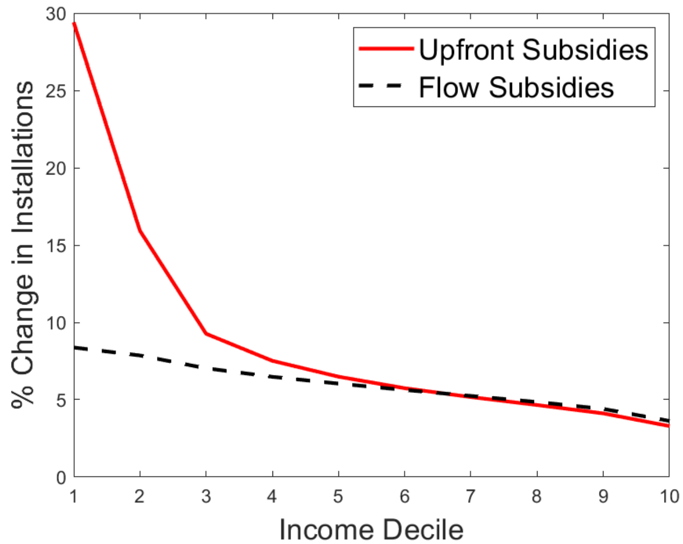


Figure 8: Upfront vs. Flow Subsidies. The lines show the percentage change in installations across income deciles associated with increasing upfront subsidies (red line) and flow subsidies (black dotted line). Both subsidy increases lead to a 10% increase in fiscal costs.

5 Robustness and Additional Results

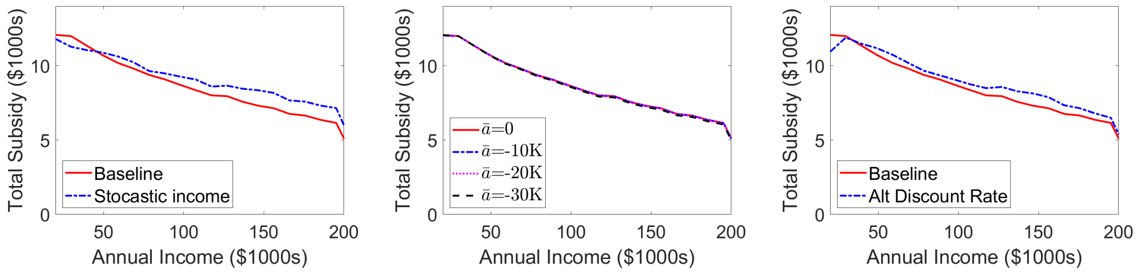
5.1 Stochastic Income

In this section, we consider a version of our model in which household income follows a linear Gaussian process with both persistent and transitory shocks, a common way of modeling stochastic earning processes in the macro and labor literature (see, e.g., Storesletten, Telmer, and Yaron (2004), Meghir and Pistaferri (2004), Guvenen (2009), Heathcote, Storesletten, and Violante (2010), Krueger, Mitman, and Perri (2016), or Guvenen et al. (2021)). We first give an overview of the model before showing how stochastic income affects our main results—additional model details and results are in Appendix C.13.

Model Let z_{it} denote the persistent component of household i's income. We assume this follows an AR(1) process as

$$z_{it} = \rho z_{it-1} + \eta_{it},$$

where ρ is a parameter that dictates the persistence of income, and η_{it} is a persistent income shock drawn from a normal distribution with mean zero and variance σ_η^2.



(a) Stochastic Income (b) Borrowing Constraints (c) Discount Rate

Figure 9: The figure on the left shows the average subsides under the production-maximizing subsidy scheme in the baseline model and a model with a stochastic income process. The figure in the middle shows the average subsides under the production-maximizing subsidy scheme under various levels of borrowing constraints. The figure on the right shows the average subsides under the production-maximizing subsidy scheme when households have a discount rate of $\beta = 0.85$. For each alternative specification, we re-estimate the model and re-solve for the production-maximizing subsidies.

Log household earnings in year t are given by

$$\log y_{it} = z_{it} + \varepsilon_{it},$$

where ε_{it} is a normally distributed transitory shock with mean zero and variance σ_ε^2. As is standard, we assume households know the distributions of ε_{it} and η_{it} but only learn the values of the two shocks in year t. Households make installation, consumption, and savings decisions to maximize lifetime expected utility by integrating over the distribution of future shocks.

We calibrate the parameters ρ, σ_ε^2 and σ_η^2 using the estimates from Storesletten, Telmer, and Yaron (2004), who estimate these parameters using earnings data from the Panel Study of Income Dynamics. We then re-estimate the remaining parameters using the indirect inference procedure described in Section 3.3. We present the estimated parameters and model fit in Appendix C.13.2. Both the estimated parameters and model fit results are very similar to those with the baseline model.

Results The production-maximizing subsidy schedules in the baseline model and the model with stochastic income are displayed in Figure 9a. The production-maximizing subsidy schedules are very similar in both models. In Appendix C.13.2, we present the optimal subsidies, distribution of installations, and production levels associated with welfare-maximizing subsidies. The optimal subsidy schemes, distribution of installations, and production levels are very similar to those under the baseline model.

5.2 Alternative Borrowing Limits

In our baseline model, we assumed that households must hold minimum assets of $\bar{a} = 0$. In this section, we examine the robustness of our results to alternative values of the minimum asset level.[32] For each alternative minimum asset level, we re-estimate the model and re-solve for the production-maximizing subsidies. The production-maximizing subsidy schedules for each level of minimum assets are displayed in Figure 9b. The production-maximizing subsidy schedules are very similar across all specifications.

5.3 Alternative Household Discount Rate

In our baseline model, we assumed that households' discount rate is given by $\beta = \frac{1}{1+r}$, where r is the market interest rate.[33] However, research by De Groote and Verboven (2019) shows that households heavily discount the future monetary benefits associated with solar panel installations.

In this section, we re-estimate the model and re-calculate optimal subsidies when households use a discount rate of $\beta = 0.85$, based on the results from De Groote and Verboven (2019). We display the optimal subsidies under this alternative discount rate are displayed in Figure 9c. The optimal subsidy schedule is similar to our baseline optimal subsidy schedule.

5.4 Maximizing Environmental Benefit

Our main results have focused on a government that maximizes residential solar production subject to a fiscal cost constraint. However, a large literature[34] highlights that the environmental benefits of solar panels vary dramatically depending on where the solar panels are installed. Therefore, the subsides which maximize solar production are not necessarily the best for the environment.

In Appendix C.11, we solve for the income-contingent subsidy schedule that maximizes environmental benefits. As in Section 4.3, we calculate environmental benefits of solar panels by estimating NERC-region-level marginal damages of electricity production during daytime hours using the estimation strategy and data from Holland et al.

[32]In all specifications, households may not have negative assets at the end of year T.

[33]We follow Sexton et al. (2021) and set $r = .05$ in our baseline model.

[34]See e.g. Siler-Evans et al. (2013), Graff Zivin, Kotchen, and Mansur (2014), Holland et al. (2016), Millstein et al. (2017), Callaway, Fowlie, and McCormick (2018), Holland et al. (2020), Brown and O'Sullivan (2020), Lamp and Samano (2023) Borenstein and Bushnell (2022), Sexton et al. (2021), and Colas and Saulnier (2023).

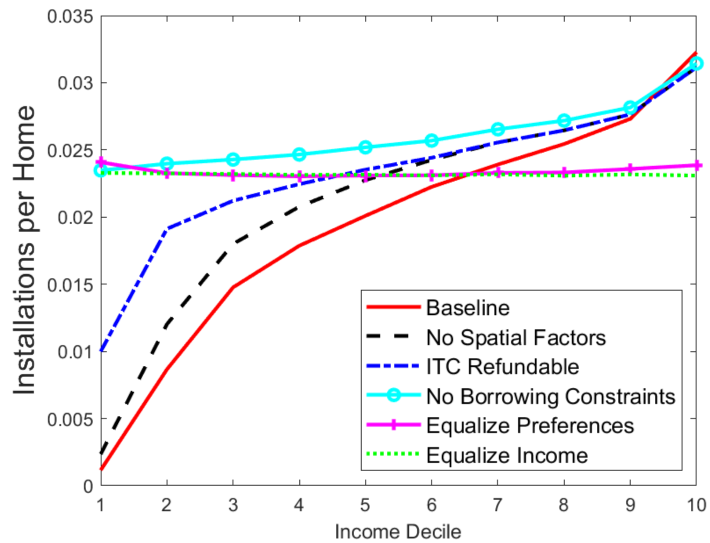


Figure 10: Model-Based Decomposition. The graph shows the average number of installations per homeowner across deciles of household income over various model specifications. See text for details on each specification.

(2020). We provide additional details in Appendix C.8. The benefit-maximizing schedule is very similar to the production-maximizing schedule and leads to a 2.3% increase in environmental benefits of residential solar nationally.

5.5 Why are Installations Increasing in Income?

Five main mechanisms in our model collectively generate a relationship between household income and installations per homeowner. First, high-income households tend to live in states with more generous subsidies and higher electricity prices.[35] Second, the Federal Investment Tax Credit is nonrefundable, and therefore, low-income households may not be able to take full advantage of this tax credit because their tax burden is too low. Third, households face borrowing constraints, and therefore, low-income households may not be able to afford solar panels despite the long-run monetary benefits. Fourth, high-income households may have stronger preferences for installations.[36] Finally, differences in income and assets imply different marginal utilities of consumption.

We now perform a model-based decomposition to understand the role played by each

[35]The tract-level correlation between average income and electricity prices is 0.18. The correlations between average income and property taxes, production subsidies, and cost subsidies are 0.11, 0.10, and 0.06, respectively.

[36]The correlation between income and tract-level college-educated share and democrat share capture these differences in preferences.

of these features in generating the positive relationship between household income and installations. Specifically, we remove each of these features one by one and re-simulate the model. All changes to the model specification are cumulative.[37] Recall that we only consider homeowners in our analysis. Therefore, we plot the relationship between income and installations per homeowner rather than household.[38]

The results are displayed in Figure 10. The solid red line shows the baseline case. Less than 0.1 percent of homeowners in the bottom income decile install panels compared to 3.6 percent of homeowners in the top decile. We begin our decomposition by removing all spatial factors and assuming all subsidies, prices, and levels of solar irradiance are drawn randomly from their respective unconditional distributions. This leads to a slight increase in installations for poorer households, as states with lower income levels tend to have less generous subsidies and lower electricity prices.

The non-refundable nature of Federal Investment Tax Credits may play an essential role in explaining low installation rates for low-income households. We next simulate a version of the model in which we additionally assume that the Federal Investment Tax Credit pays in full in the year of installation. This change leads to a large increase in installations for lower income levels but no change in installations for households with higher income, who already have a large enough tax burden to receive the full credit in the year of installation.

Next, we simulate a version of the model in which households additionally do not face a borrowing constraint and can borrow freely against future income. This change leads to a large increase in installations for poorer households but no change for higher-income households, who do not need to borrow to finance solar panels. We next remove the correlation between household income and preferences by setting X_ℓ^{Coll} and X_ℓ^{Pol} to the median values in the data. This equalization lowers installation rates for high-income households. Finally, we remove the role of income directly by setting all household income levels to the national mean. This fully removes the relationship between income and installations.

In summary, this decomposition suggests that borrowing constraints and the non-refundability of the Federal Investment Tax Credit play major roles in generating the positive relationship between installations and income. Preferences and spatial differ-

[37]In Appendix C.9, we perform these modifications individually, rather than cumulatively.

[38]We show the same graph with installations per household, rather than homeowner, in Appendix C.9. Homeownership rates are strongly increasing in household income. Therefore, the installation rates per household, including non-homeowners, increase more strongly in income than those shown here.

ences between prices, subsidies, and sunlight play relatively minor roles.

6 Conclusion

We study the optimal design of income-contingent subsidies for residential solar panels. We show robust reduced-form evidence that the partial-elasticity of solar production with respect to subsidies is decreasing in income, suggesting that means-tested subsidies could induce greater solar production per dollar of public funds compared to income-neutral subsidies. Simulations from a quantitative model reveal that optimally set income-contingent subsidies lead to a much more equitable distribution of public funds and an increase in residential solar production. Therefore, means-tested solar subsidies are justified on both equity and cost-efficiency grounds.

Future work could extend the empirical exercise here to other green products, such as energy-efficient appliances or heat pumps. The equity-efficiency trade-offs associated with subsidies for heat pumps are likely very different than those for solar panels subsidies, as heat pump adoption rates do not correlate strongly with income (Davis, 2023). It would also be interesting to consider a government that can offer financing programs for solar panels in addition to means-tested subsidies. These financing programs would alleviate borrowing constraints for low-income households, which would change the benefits of providing income-contingent subsidies to those households. We leave these questions for future research.

References

Anderson, Soren T, Ryan Kellogg, and James M Sallee. 2013. "What do consumers believe about future gasoline prices?" *Journal of Environmental Economics and Management*, 66(3): 383–403.

Battistin, Erich, Richard Blundell, and Arthur Lewbel. 2009. "Why is consumption more log normal than income? Gibrat's law revisited." *Journal of Political Economy*, 117(6): 1140–1154.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A unified framework for measuring preferences for schools and neighborhoods." *Journal of political economy*, 115(4): 588–638.

Bento, Antonio M, Lawrence H Goulder, Mark R Jacobsen, and Roger H Von Haefen. 2009. "Distributional and efficiency impacts of increased US gasoline taxes." *American Economic Review*, 99(3): 667–699.

Borenstein, Severin. 2012. "The redistributional impact of nonlinear electricity pricing." *American Economic Journal: Economic Policy*, 4(3): 56–90.

Borenstein, Severin. 2017. "Private net benefits of residential solar PV: The role of electricity tariffs, tax incentives, and rebates." *Journal of the Association of Environmental and Resource Economists*, 4(S1): S85–S122.

Borenstein, Severin, and James B. Bushnell. 2022. "Do Two Electricity Pricing Wrongs Make a Right? Cost Recovery, Externalities, and Efficiency." *American Economic Journal: Economic Policy*, 14(4): 80–110.

Borenstein, Severin, and Lucas W Davis. 2016. "The distributional effects of US clean energy tax credits." *Tax Policy and the Economy*, 30(1): 191–234.

Brown, Patrick R, and Francis M O'Sullivan. 2020. "Spatial and temporal variation in the value of solar power across United States electricity markets." *Renewable and Sustainable Energy Reviews*, 121: 109594.

Burr, Chrystie. 2014. "Subsidies, tariffs and investments in the solar power market." *University of Colorado-Boulder Working Paper*.

Cahana, Michael, Natalia Fabra, Mar Reguant, and Jingyuan Wang. 2022. "The distributional impacts of real-time pricing."

Callaway, Duncan S, Meredith Fowlie, and Gavin McCormick. 2018. "Location, location, location: The variable value of renewable energy and demand-side efficiency resources." *Journal of the Association of Environmental and Resource Economists*, 5(1): 39–75.

Colas, Mark, and Emmett Saulnier. 2023. "Optimal Solar Subsidies."

Colas, Mark, Sebastian Findeisen, and Dominik Sachs. 2021. "Optimal need-based financial aid." *Journal of Political Economy*, 129(2): 492–533.

Crago, Christine Lasco, and Ilya Chernyakhovskiy. 2017. "Are policy incentives for solar power effective? Evidence from residential installations in the Northeast." *Journal of Environmental Economics and Management*, 81: 132–151.

Dauwalter, Travis E, and Robert I Harris. 2023. "Distributional Benefits of Rooftop Solar Capacity." *Journal of the Association of Environmental and Resource Economists*, 10(2): 487–523.

Davis, Lucas W. 2023. "The Economic Determinants of Heat Pump Adoption." National Bureau of Economic Research.

Davis, Lucas W, and Christopher R Knittel. 2019. "Are fuel economy standards regressive?" *Journal of the Association of Environmental and Resource Economists*, 6(S1): S37–S63.

De Groote, Olivier, and Frank Verboven. 2019. "Subsidies and time discounting in new technology adoption: Evidence from solar photovoltaic systems." *American Economic Review*, 109(6): 2137–72.

Dorsey, Jackson, and Derek Wolfson. 2023. "Unequal Uptake: Assessing Distributional Disparities in the Residential Solar Market."

Feger, Fabian, Nicola Pavanini, and Doina Radulescu. 2022. "Welfare and redistribution in residential electricity markets with solar power." *The Review of Economic Studies*, 89(6): 3267–3302.

Fried, Stephie, Kevin Novan, and William B Peterman. 2018. "The distributional effects of a carbon tax on current and future generations." *Review of Economic Dynamics*, 30: 30–46.

Fried, Stephie, Kevin Novan, and William B Peterman. 2022. "Understanding the Inequality and Welfare Impacts of Carbon Tax Policies." Working Paper.

Gillingham, Kenneth, and Tsvetan Tsvetanov. 2019. "Hurdles and steps: Estimating demand for solar photovoltaics." *Quantitative Economics*, 10(1): 275–310.

Goulder, Lawrence H, Marc AC Hafstead, GyuRim Kim, and Xianling Long. 2019. "Impacts of a carbon tax across US household income groups: What are the equity-efficiency trade-offs?" *Journal of Public Economics*, 175: 44–64.

Graff Zivin, Joshua S, Matthew J Kotchen, and Erin T Mansur. 2014. "Spatial and temporal heterogeneity of marginal emissions: Implications for electric cars and other electricity-shifting policies." *Journal of Economic Behavior & Organization*, 107: 248–268.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura. 2014. "Income taxation of US households: Facts and parametric estimates." *Review of Economic Dynamics*, 17(4): 559–581.

Guvenen, Fatih. 2009. "An empirical investigation of labor income processes." *Review of Economic dynamics*, 12(1): 58–79.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song. 2021. "What do data on millions of US workers reveal about lifecycle earnings dynamics?" *Econometrica*, 89(5): 2303–2339.

Hahn, Robert W, and Robert D Metcalfe. 2021. "Efficiency and equity impacts of energy subsidies." *American Economic Review*, 111(5): 1658–1688.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante. 2010. "The macroeconomic implications of rising wage inequality in the United States." *Journal of political economy*, 118(4): 681–722.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante. 2017. "Optimal tax progressivity: An analytical framework." *The Quarterly Journal of Economics*, 132(4): 1693–1754.

Holland, Stephen P, Erin T Mansur, Nicholas Z Muller, and Andrew J Yates. 2016. "Are there environmental benefits from driving electric vehicles? The importance of local factors." *American Economic Review*, 106(12): 3700–3729.

Holland, Stephen P, Erin T Mansur, Nicholas Z Muller, and Andrew J Yates. 2019. "Distributional effects of air pollution from electric vehicle adoption." *Journal of the Association of Environmental and Resource Economists*, 6(S1): S65–S94.

Holland, Stephen P, Erin T Mansur, Nicholas Z Muller, and Andrew J Yates. 2020. "Decompositions and policy consequences of an extraordinary decline in air pollution from electricity generation." *American Economic Journal: Economic Policy*, 12(4): 244–274.

Hughes, Jonathan E, and Molly Podolefsky. 2015. "Getting green with solar subsidies: evidence from the California solar initiative." *Journal of the Association of Environmental and Resource Economists*, 2(2): 235–275.

Ito, Koichiro. 2014. "Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing." *American Economic Review*, 104(2): 537–563.

Jacobsen, Mark R. 2013. "Evaluating US fuel economy standards in a model with producer and household heterogeneity." *American Economic Journal: Economic Policy*, 5(2): 148–187.

Jordan, Dirk C, and Sarah R Kurtz. 2013. "Photovoltaic degradation rates—an analytical review." *Progress in photovoltaics: Research and Applications*, 21(1): 12–29.

Kiribrahim-Sarikaya, Ozgen, and Yueming Qiu. 2023. "Distributional Effects of Residential Solar Subsidies."

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri. 2016. "Macroeconomics and household heterogeneity." In *Handbook of macroeconomics*. Vol. 2, 843–921. Elsevier.

Lamp, Stefan, and Mario Samano. 2023. "(Mis)allocation of Renewable Energy Sources." *Journal of the Association of Environmental and Resource Economists*, 10(1): 195–229.

Langer, Ashley, and Derek Lemoine. 2022. "Designing dynamic subsidies to spur adoption of new technologies." *Journal of the Association of Environmental and Resource Economists*, 9(6): 1197–1234.

Linn, Joshua. 2022. "Is there a trade-off between equity and effectiveness for electric vehicle subsidies." *Resources for the Future.*

Meghir, Costas, and Luigi Pistaferri. 2004. "Income variance dynamics and heterogeneity." *Econometrica*, 72(1): 1–32.

Millstein, Dev, Ryan Wiser, Mark Bolinger, and Galen Barbose. 2017. "The climate and air-quality benefits of wind and solar power in the United States." *Nature Energy*, 2(9): 1–10.

Morehouse, John M. 2021. "Carbon Taxes in Spatial Equilibrium." *Working Paper.*

Reguant, Mar. 2019. "The efficiency and sectoral distributional impacts of large-scale renewable energy policies." *Journal of the Association of Environmental and Resource Economists*, 6(S1): S129–S168.

Sexton, Steven, A Justin Kirkpatrick, Robert I Harris, and Nicholas Z Muller. 2021. "Heterogeneous solar capacity benefits, appropriability, and the costs of suboptimal siting." *Journal of the Association of Environmental and Resource Economists*, 8(6): 1209–1244.

Siler-Evans, Kyle, Inês Lima Azevedo, M Granger Morgan, and Jay Apt. 2013. "Regional variations in the health, environmental, and climate benefits of wind and solar generation." *Proceedings of the National Academy of Sciences*, 110(29): 11768–11773.

Snashall-Woodhams, Nicholas. 2019. "Targeting Solar Subsidies." Working Paper.

Storesletten, Kjetil, Christopher I Telmer, and Amir Yaron. 2004. "Consumption and risk sharing over the life cycle." *Journal of monetary Economics*, 51(3): 609–633.

Yu, Jiafan, Zhecheng Wang, Arun Majumdar, and Ram Rajagopal. 2018. "DeepSolar: A machine learning framework to efficiently construct a solar deployment database in the United States." *Joule*, 2(12): 2605–2617.

A Theoretical Appendix

A.1 Proof of Proposition 1

Consider a variation of s denoted by δs. We are interested in subsidy variations that are cost-neutral, such that

$$dCost = \int_{\underline{y}}^{\overline{y}} \frac{\delta Cost}{\delta s\left(y\right)} \delta s\left(y\right) dy = \int_{\underline{y}}^{\overline{y}} \delta s\left(y\right) \left(\frac{\partial K_y}{\partial s_y} \bar{s} + K_y(\bar{s}) \right) dy = 0 \tag{12}$$

and that are progressive, meaning that

$$\left(y'' - y'\right) \left(\delta s(y'') - \delta s(y')\right) < 0 \tag{13}$$

for $y'' \neq y'$. The total change in production associated with the variation δs is equal to

$$dProd = \int_{\underline{y}}^{\overline{y}} \frac{\partial K_y}{\partial s_y} \delta s\left(y\right) dy.$$

We can think of the change in cost associated with δs as the costs from additional households, who each receive $\bar{s}$, plus the costs from non-additional households, who each receive an additional $\delta s(y)$. Explicitly, we can rewrite $dCost$ as the current subsidy level times the change in production multiplied by a constant:

$$dCost = \bar{s} \times dProd + A, \tag{14}$$

where we define

$$\int_{\underline{y}}^{\overline{y}} \delta s\left(y\right) K_y(\bar{s}) dy \equiv A.$$

Intuitively, A gives the change in cost associated with non-additional households, while $\bar{s} \times dProd$ gives the change in cost associated with additional households. The change in production will be large relative to cost when the cost to non-additional households, given by A, is small, all else equal. We can rewrite A in terms of cost-effectiveness as

$$A = \int_{\underline{y}}^{\overline{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(y) dy.$$

Since δs is decreasing in income, and since δs is a cost-neutral variation, there must exist some "cutoff" income level $\tilde{y}$ such that $\delta s(y) \geq 0$ for $y \leq \tilde{y}$ and $\delta s(y) \leq 0$ for

$y > \tilde{y}$. We can rewrite A as an integral over households with income below $\tilde{y}$ (who receive subsidy increases), plus an integral over households with income above $\tilde{y}$ (who receive subsidy reductions):

$$A = \int_{\underline{y}}^{\tilde{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(y) dy + \int_{\tilde{y}}^{\overline{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(y) dy. \tag{15}$$

Let $\eta(\tilde{y})$ denote the cost-effectiveness associated with this cutoff income level $\tilde{y}$.

We now consider two cases. First, we consider the case in which η is decreasing in income. Second, we consider the case in which η is increasing in income.

Case 1: η Decreasing In Income Assume η is weakly decreasing in income and $\eta(\underline{y}) > \eta(\overline{y})$. Thus, it must be the case that $\eta^{-1}(y) \leq \eta^{-1}(\tilde{y})$ for $y \leq \tilde{y}$ and $\eta^{-1}(y) \geq \eta^{-1}(\tilde{y})$ for $y > \tilde{y}$ with at least one of these two inequalities holding strictly. As $\delta s\left(y\right) \geq 0$ for $y \leq \tilde{y}$, it must be that

$$\int_{\underline{y}}^{\tilde{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(y) dy \leq \int_{\underline{y}}^{\tilde{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(\tilde{y}) dy. \tag{16}$$

Further, since $\delta s\left(y\right) < 0$ for $y > \tilde{y}$, we know that

$$\int_{\tilde{y}}^{\overline{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(y) dy \leq \int_{\tilde{y}}^{\overline{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(\tilde{y}) dy. \tag{17}$$

At least one of the above inequalities must hold strictly, because either $\eta^{-1}(y) > \eta^{-1}(\tilde{y})$ for $y \leq \tilde{y}$ or $\eta^{-1}(y) < \eta^{-1}(\tilde{y})$ for $y > \tilde{y}$.

Using these inequalities in equations (16) and (17), we can rewrite equation (15) as

$$\begin{aligned} A = \int_{\underline{y}}^{\tilde{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(y) dy + \int_{\tilde{y}}^{\overline{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(y) dy < \\ \int_{\underline{y}}^{\tilde{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(\tilde{y}) dy + \int_{\tilde{y}}^{\overline{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(\tilde{y}) dy. \end{aligned} \tag{18}$$

We can rewrite this inequality as

$$A < \eta^{-1}(\tilde{y}) \left(\int_{\underline{y}}^{\overline{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} dy \right) = \eta^{-1}(\tilde{y}) \times dProd.$$

Adding $\bar{s} \times dProd$, and noting the cost-neutrality constrant, yields

$$\underbrace{\bar{s} \times dProd + A}_{=dCost=0} < \left(\bar{s} + \eta^{-1}(\tilde{y})\right) dProd.$$

Finally, dividing both sides by $(\bar{s} + \eta^{-1}(\tilde{y})) > 0$ yields

$$dProd > 0.$$

Case 2: η Increasing In Income Assume η is weakly increasing in income and $\eta(\underline{y}) < \eta(\overline{y})$. Thus, it must be the case that $\eta^{-1}(y) \geq \eta^{-1}(\tilde{y})$ for $y \leq \tilde{y}$ and $\eta^{-1}(y) \leq \eta^{-1}(\tilde{y})$ for $y > \tilde{y}$ with one of these two inequalities holding strictly. Following the logic from the Case 1, we can then show that

$$A = \int_{\underline{y}}^{\tilde{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(y) dy + \int_{\tilde{y}}^{\overline{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(y) dy > \int_{\underline{y}}^{\tilde{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(\tilde{y}) dy + \int_{\tilde{y}}^{\overline{y}} \delta s\left(y\right) \frac{\partial K_y}{\partial s_y} \eta^{-1}(\tilde{y}) dy. \quad (19)$$

Following the same algebriac steps as the previous case yields

$$\underbrace{\bar{s} \times dProd + A}_{=sCost=0} > \left(\bar{s} + \eta^{-1}(\tilde{y})\right) dProd$$

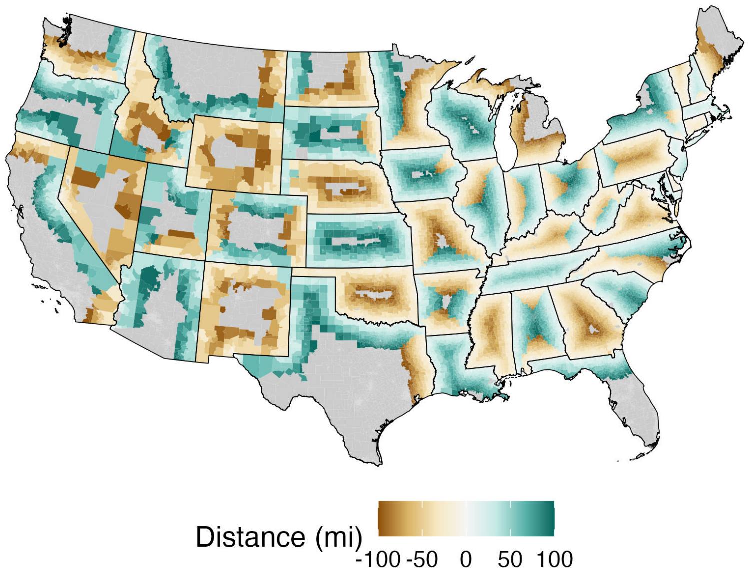
which implies that $dProd < 0$ since $(\bar{s} + \eta^{-1}(\tilde{y})) > 0$.

B Reduced-Form Appendix

B.1 Non-Reweighted Border Discontinuities in Subsidy Generosity and Log Production per Capita

Here we present border discontinuity graphs for subsidy generosity and log production per capita where we do not reweight observations as in our main specification. As before, we define a tract's location to the nearest border as the positive distance to the border for tracts on the side of the border with more generous subsidies and the negative distance to the border for tracts on the side with less generous subsidies. Figure A1 shows the distance to nearest border for each census tract, with the less-generous side

Figure A1: Distances to nearest border by census tract.

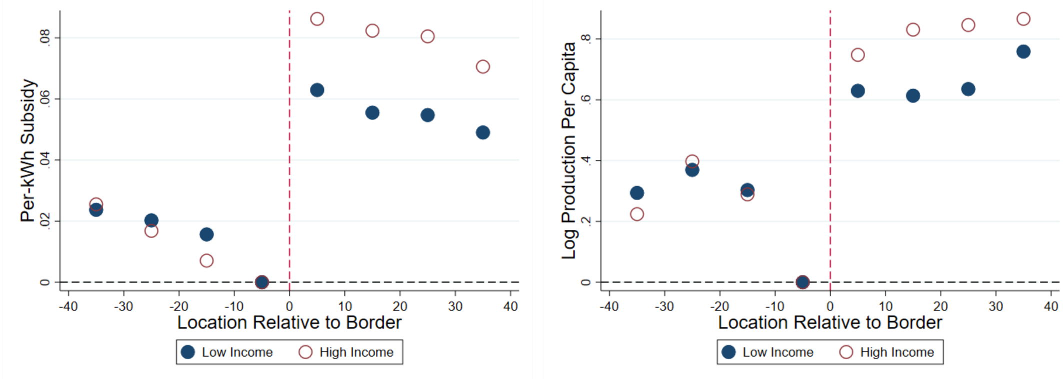


of the border in brown and more generous side in green. We categorize tracts into 10-mile-wide bins based on this location relative to the border and regress the variable in questions (either subsidy generosity or log production per capita) on state-border fixed effects and fixed effects for these location bins. We run these regressions separately for high- and low-income tracts.

Figure A2a plots the estimated location-bin fixed effects for a regression on subsidy generosity for high-income and low-income tracts. Positive values on the X-axis represent tracts on the side of the border with more generous subsidies, and negative values on the X-axis represent tracts on the side with less generous subsidies.[39] Mechanically, subsidy generosity for both groups increase sharply as we move to the side with more generous subsidies. The increase in subsidy generosity is larger for high-income tracts revealing that high income tracts tend to be located around borders with larger differences in subsidy generosity. Figure A2a plots the estimated location-bin fixed effects for a regression on log production per capita for high-income and low-income tracts.

Taken together, the two graphs show that crossing borders from a state with less generous subsidies to a state with more generous subsidies is associated with both larger subsidy increases and larger production rate increases for high income tracts. The reweighting procedure we utilize in Section 3.2 allows us to compare discontinuities

[39]The regressions omit the location-bin fixed effect for the location bin nearest to the border on the less generous subsidy side. We can therefore interpret these estimated location bin fixed effects as the conditional average of subsidy generosity in a given location bin relative to this omitted bin.



(a) Subsidy Generosity (b) Log Production Per Capita

Figure A2: Border Discontinuities in Subsidy Generosity and Log Production per Capita without Reweighting. The graph plots estimated location-bin fixed effects from a regression of subsidy generosity (Panel (a)) or log production per capita (Panel (b)) separately for high-income (orange) tracts and low-income (green) tracts on border fixed effects and location-bin fixed effects. Positive values on the X-axis represent households on the side of the border with more generous subsidies, and negative values on the X-axis indicate the side of the border with less generous subsidies.

in log production across borders with consistent weights across low-income and high-income regressions.

B.2 Border Discontinuities in Demographics and Household Income

In this appendix, we plot average tract-level demographics and average income around state borders. We categorize tracts into 10-mile-wide bins based on their location relative to the border and regress various demographics on state-border fixed effects and fixed effects for these location bins. Figure A3 plots these estimated location-bin fixed effects for percent with average household income, college degree, and percent democrat, respectively. There is no discontinuity in these characteristics as we cross to the side of the border with more generous subsidies.

B.3 Full coefficient tables

Table A1 reports all coefficients for the specifications in Table 1.

B.4 Linear specification without border discontinuities

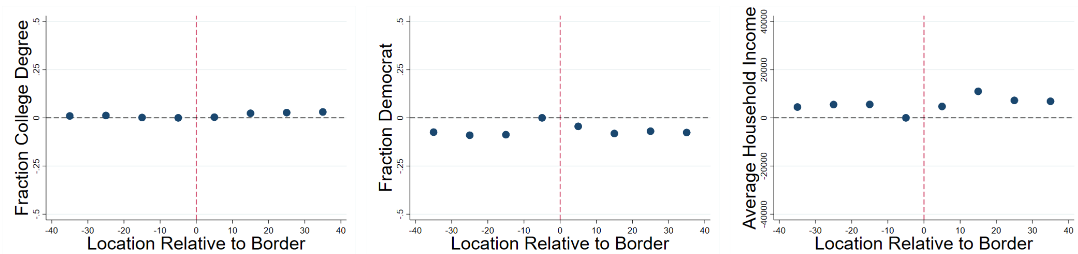
Here, we estimate the specification where cost-effectiveness varies linearly in log income using all census tracts. The first two columns of Table A2 show the results. The

Table A1: Effect of Subsidies on Log Production per Capita

Border Polynomial Deg.	0		3		5	
Bandwidth	40 mi	80 mi	40 mi	80 mi	40 mi	80 mi
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	-3.40*	-4.13***	-4.53***	-2.53*	-3.53***	-3.33***
	(1.96)	(1.53)	(1.19)	(1.28)	(1.09)	(1.14)
Subsidy	6.04***	6.53***	4.31**	4.74**	3.57*	4.02**
	(0.949)	(1.08)	(2.01)	(1.93)	(2.11)	(1.95)
Log Income	-0.323	-0.371	-0.173	-0.239	-0.226	-0.274
	(0.228)	(0.252)	(0.172)	(0.249)	(0.177)	(0.208)
Elec. Price	-39.4	-43.9	-0.077	-16.8	-12.9	-16.9
	(27.1)	(31.3)	(26.1)	(34.5)	(25.2)	(25.9)
Solar Irradiance	1.26***	1.37***	0.988***	0.728***	0.844***	0.771***
	(0.415)	(0.326)	(0.245)	(0.207)	(0.237)	(0.223)
Percent College	0.877***	1.02***	0.965***	1.10***	0.954***	1.05***
	(0.175)	(0.157)	(0.164)	(0.140)	(0.167)	(0.136)
Percent Owner	0.329	0.267	0.222	0.170	0.212	0.116
	(0.346)	(0.259)	(0.333)	(0.250)	(0.337)	(0.247)
Percent Democrat	0.334	0.543*	0.213	0.444*	0.234	0.373
	(0.370)	(0.273)	(0.391)	(0.251)	(0.376)	(0.258)
Population Density	-5.23***	-5.27***	-5.57***	-5.43***	-5.65***	-5.64***
	(0.449)	(0.440)	(0.496)	(0.452)	(0.496)	(0.449)
Population Density sq	2.17***	2.15***	2.33***	2.26***	2.36***	2.33***
	(0.257)	(0.242)	(0.286)	(0.253)	(0.277)	(0.245)
Subsidy × Log Income	-1.50**	-1.92**	-1.76***	-1.92**	-2.02***	-2.07***
	(0.704)	(0.913)	(0.428)	(0.845)	(0.450)	(0.681)
Elec. Price × Log Income	3.96	4.59	2.68	3.32	3.52**	3.99**
	(2.52)	(2.82)	(1.66)	(2.55)	(1.68)	(1.98)
Fit statistics						
Observations	20,187	30,410	20,187	30,410	20,187	30,410
R^2	0.48	0.49	0.54	0.55	0.19	0.56

Clustered (State) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Estimates of coefficients from Equation (4). Sample is limted to tracts within either 40 or 80 miles bandwidths to state borders. All regressions include controls for tract-level income, electricity prices, electricity prices interacted with income, solar irradiance, population density and population density squared, a battery of tract-level demographic measures, and border-specific polynomials in location relative to border.



(a) Percent College Degree (b) Percent Democrat (c) Average Household Income

Figure A3: Border Discontinuities in Demographics and Household Income. Each graph plots estimated location-bin fixed effects from a regression of the variable in question on border fixed effects and location-bin fixed effects. Positive values on the X-axis represent households on the side of the border with more generous subsidies, and negative values on the X-axis indicate the side of the border with less generous subsidies.

second column, with division fixed effects, suggests that a 1 cent per KWh increase in subsidies leads to a nearly 6.5 percent increase in solar production per capita. The following two columns of Table A2 have estimates of equation (3), where we specify that cost-effectiveness is linear in log income. We have again "de-medianed" log income by subtracting the median log income level from our log income variable. Using estimates from column 4, a 1 cent per-kWh increase in subsidies is associated with a 7.3% increase in solar production per capita for a tract at the median income level, and, consistent with our border discontinuity result, is decreasing in income. The same 1 cent per-kWh subsidy increase is associated with only a 5.1'% increase in solar production per capita for a tract at the 90th percentile of income distribution. The empirical partial elasticity of solar production with respect to subsidies for tracts at the median income level is 44 percent higher than the partial elasticity for tracts at the 90th percentile income level—somewhat larger than our border discontinuity result.

B.5 Estimating cost-effectiveness with zero production tracts

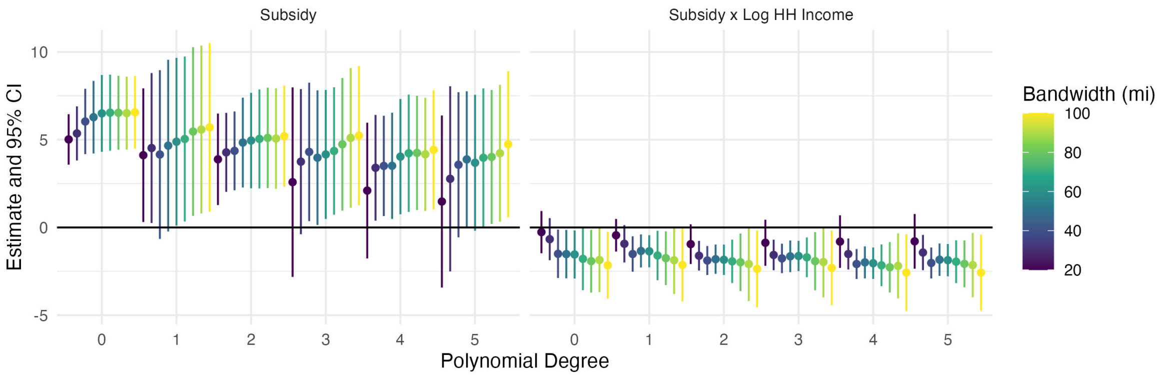
Our main specification only includes tracts with positive solar production since it requires taking the log of production as the outcome variable. However, about a quarter of all census tracts have no residential solar installations in the Deepsolar data, representing about 21 percent of the US population. We use a strategy similar to that of Gillingham and Tsvetanov (2019) to account for these zero-production tracts—estimating a logit model of the effect of subsidies on the probability of a tract having positive residential solar production. We can then use the logit model in conjunction with our main results as a "hurdle model." The hurdle model combines an "extensive margin" partial

Table A2: Effect of Subsidies on Log Production per Capita

Model:	(1)	(2)	(3)	(4)
Variables				
Subsidy	4.67***	6.28***	5.51***	7.34***
	(1.49)	(1.22)	(1.70)	(1.33)
Log Income	-0.250	-0.405	-0.208	-0.356
	(0.443)	(0.435)	(0.342)	(0.350)
Population Density	-5.46***	-5.47***	-5.50***	-5.52***
	(0.522)	(0.416)	(0.528)	(0.433)
Population Density sq	2.30***	2.32***	2.32***	2.34***
	(0.374)	(0.325)	(0.373)	(0.330)
Percent College	0.384	0.463*	0.430	0.511*
	(0.297)	(0.258)	(0.293)	(0.260)
Percent Owner	0.441	0.470	0.460	0.489
	(0.334)	(0.330)	(0.340)	(0.337)
Percent Democrat	0.983***	0.790***	0.951***	0.750**
	(0.256)	(0.278)	(0.258)	(0.282)
Elec. Price	-58.0	-71.9**	-76.7**	-93.9***
	(39.4)	(34.3)	(33.7)	(25.8)
Solar Irradiance	1.44***	1.52***	1.45***	1.53***
	(0.256)	(0.304)	(0.259)	(0.304)
Elec. Price × Log Income	6.20*	7.18**	7.87***	9.12***
	(3.49)	(3.21)	(2.91)	(2.39)
Subsidy × Log Income			-3.69***	-4.23***
			(1.19)	(1.02)
Fixed-effects				
Region	Yes		Yes	
Division		Yes		Yes
Fit statistics				
Observations	49,010	49,010	49,010	49,010
R^2	0.57	0.58	0.57	0.58
Within R^2	0.35	0.34	0.35	0.34

Clustered (State) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Figure A4: Robustness of border discontinuity estimates to different bandwidths and polynomial degrees.

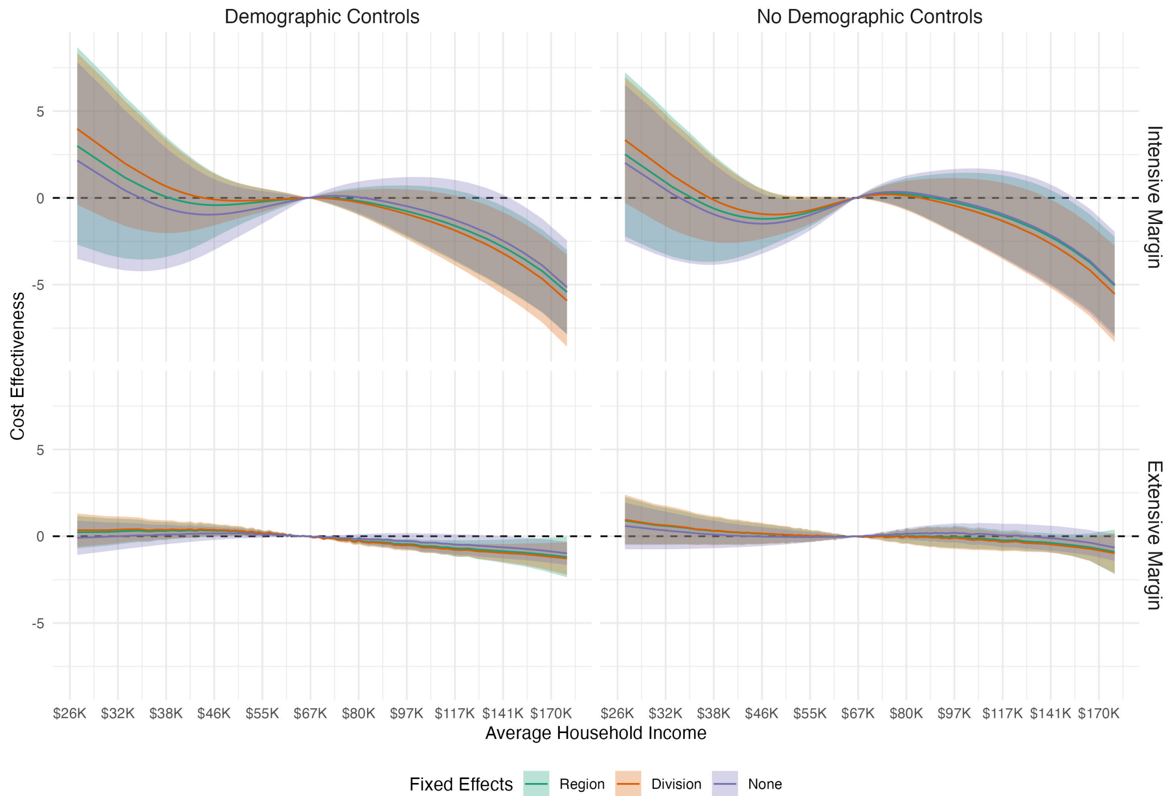


elasticity—the effect of a marginal subsidy increase on the probability of a tract having positive solar production—with an "intensive margin" partial elasticity—the effect of a marginal subsidy increase on log production per capita. These intensive margin elasticities are what we present in the main body of the paper. In the hurdle model, our cost-effectiveness measure is simply the sum of the extensive and intensive margin partial elasticities. We estimate all of the same specifications for the extensive margin as we do for the intensive margin. To calculate the extensive margin partial elasticity from the logit model, we take the average marginal effect of subsidies for different income groups on the probability of having positive solar production, divided by the total number of tracts with positive production. Section B.6 shows extensive margin results from the logit model alongside of the intensive margin results—consistently demonstrating that the magnitude of extensive margin cost-effectiveness is very small relative to that of the intensive margin.

B.6 Robustness

Border Discontinuity Bandwith and Polynomial Degree In our main specification, we use a border discontinuity with a bandwidth of 40 miles, with border fixed effects interacted with 3rd degree polynomials in distance from the border. Figure A4 shows estimates using bandwidths between 20 and 100 miles and polynomials ranging from degree 0 (just border fixed effects) to 5. The results are similar for all but the smallest bandwidths.

Figure A5: Cost effectiveness using production for different controls and fixed effects.



Alternative Controls and Fixed Effects for Nonlinear Specifications Our main results include demographic controls and census division fixed effects. Figures A5 shows robustness to alternative fixed effects and omitting the demographic controls. We show just the income spline specification for simplicity, but results are similar across different means of allowing the marginal effect of subsidies to vary by income level. Adding controls and more refined fixed effects makes our results marginally stronger. Across all specifications, the estimated cost effectiveness from the logit model is very small relative to that from the OLS model.

Controls for Other State Solar Policies A threat to identification in our border discontinuity models is a change in other policies across state borders that affects household solar installation decisions. We gather data on other state-level policies related to solar panels from DSIRE that are not monetary subsidies. We then count the number of programs in each state split into three categories: financing, access rules, and building incentives.[40] In addition to using the counts of policies, we also run models using indicators for the presence of these policies.

Tables A3, A4, and A5 show results using a 40 mile bandwith and 0, 3rd, and 5th degree border polynomials, respectively. The results are similar to our main models. The coefficient on the interaction between subsidies and log income remains very stable across all of the DSIRE specifications. The coefficient on subsidies—reflecting the partial elasticity of installations with respect to subsidies for a tract with median log income—does increase in some specifications. However, these partial elasticities are never statistically significantly different at the 5 percent level from that of our baseline model without DSIRE controls.

Alternative Outcomes We derive cost-effectiveness using production per capita as the outcome of interest. Here, we explore robustness to different outcome variables that we could have also used to derive cost-effectiveness—counts of residential solar installations and residential solar panels. Table A6 shows results for the same specifications as Table 1, but with log panels per capita as the outcome. Figure A6 shows the robustenss of these border discontinuity results for these alternative outcomes across all bandwidths and polynomial degrees. Figures A7 and A8 demonstrate that the partial elasticity of subsidies with respect to either of these alternative outcomes is also decreasing in income when we use the nonlinear specifications and all census tracts.

Alternative Models Since installations are a count variable, we estimate a poisson model limiting to tracts that only have positive installations. Figure A9 shows results using evenly spaced bins by log income. Our results are qualitatively unchanged.

[40]Financing are policies with labels of "Loan Program" and "PACE Financing" in DSIRE, access are solar policies with labels of "Solar/Wind Access Policy" and "Solar/Wind Permitting Standards", and building incentives are policies with the label of "Building Energy Codes" and "Green Building Incentives"

Table A3: Effect on Log production per capita in border discontinuity model with additional DSIRE solar policy controls. These estimates are from Equation (4) with Degree 0 border-specific polynomials and a 40 mile bandwidth. All regressions include controls for tract-level income, electricity prices, electricity prices interacted with income, solar irradiance, population density and population density squared, and a battery of tract-level demographic measures.

Model:	(1)	(2)	(3)
Variables			
Subsidy	6.04***	7.52***	6.99***
	(0.949)	(1.39)	(1.51)
Subsidy × Log Income	-1.50**	-1.44*	-1.42*
	(0.704)	(0.717)	(0.760)
Count DSIRE financing		0.262***	
		(0.065)	
Count DSIRE access		0.062	
		(0.088)	
Count DSIRE building		-0.074	
		(0.127)	
Has DSIRE financing			0.301***
			(0.092)
Has DSIRE access			0.097
			(0.127)
Has DSIRE building			-0.049
			(0.196)
Fit statistics			
Observations	20,187	20,187	20,187
R^2	0.48	0.49	0.49
Adjusted R^2	0.48	0.49	0.49

Clustered (State) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A4: Effect on Log production per capita in border discontinuity model with additional DSIRE solar policy controls. These estimates are from Equation (4) with 3rd degree border-specific polynomials and a 40 mile bandwidth. All regressions include controls for tract-level income, electricity prices, electricity prices interacted with income, solar irradiance, population density and population density squared, and a battery of tract-level demographic measures.

Model:	(1)	(2)	(3)
Variables			
Subsidy	4.31**	6.35***	6.78***
	(2.01)	(2.30)	(2.15)
Subsidy × Log Income	-1.76***	-1.82***	-1.83***
	(0.428)	(0.425)	(0.423)
Count DSIRE financing		0.126	
		(0.130)	
Count DSIRE access		-0.088	
		(0.111)	
Count DSIRE building		-0.091	
		(0.218)	
Has DSIRE financing			0.231
			(0.194)
Has DSIRE access			-0.125
			(0.136)
Has DSIRE building			-0.257
			(0.238)
Fit statistics			
Observations	20,187	20,187	20,187
R^2	0.55	0.55	0.55
Adjusted R^2	0.54	0.54	0.54

Clustered (State) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A5: Effect on Log production per capita in border discontinuity model with additional DSIRE solar policy controls. These estimates are from Equation (4) with 5th degree border-specific polynomials and a 40 mile bandwidth. All regressions include controls for tract-level income, electricity prices, electricity prices interacted with income, solar irradiance, population density and population density squared, and a battery of tract-level demographic measures.

Model:	(1)	(2)	(3)
Variables			
Subsidy	3.57^{*}	7.85^{***}	7.98^{***}
	(2.11)	(2.53)	(2.55)
Subsidy × Log Income	-2.02^{***}	-2.07^{***}	-2.08^{***}
	(0.450)	(0.449)	(0.448)
Count DSIRE financing		0.207	
		(0.147)	
Count DSIRE access		-0.185^{*}	
		(0.107)	
Count DSIRE building		-0.212	
		(0.142)	
Has DSIRE financing			0.314
			(0.195)
Has DSIRE access			-0.189
			(0.167)
Has DSIRE building			-0.482^{*}
			(0.246)
Fit statistics			
Observations	20,187	20,187	20,187
R^2	0.54	0.55	0.55
Adjusted R^2	0.52	0.53	0.53

Clustered (State) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A6: Effect of Subsidies on Log Panels per Capita

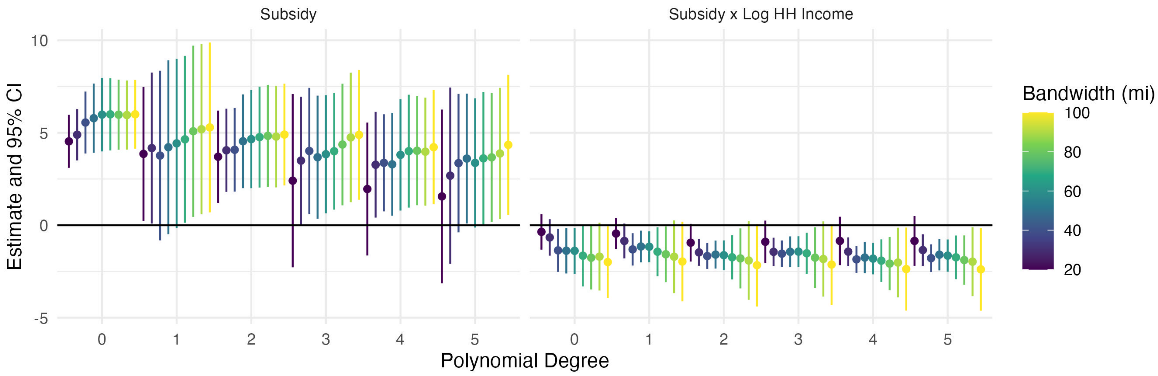
Border Polynomial Deg.	0		3		5	
Bandwidth (mi)	40 mi	80 mi	40 mi	80 mi	40 mi	80 mi
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Subsidy	6.04***	6.53***	4.31**	4.74**	3.57*	4.02**
	(0.949)	(1.07)	(2.01)	(1.93)	(2.11)	(1.95)
Subsidy × Log Income	-1.50**	-1.92**	-1.76***	-1.92**	-2.02***	-2.07***
	(0.703)	(0.912)	(0.428)	(0.845)	(0.450)	(0.681)
Fit statistics						
Observations	20,187	30,410	20,187	30,410	20,187	30,410
R^2	0.48	0.48	0.54	0.54		0.55

Clustered (State) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Estimates of coefficients from Equation (4), but with log panels per capita as the outcome. Sample is limted to tracts within either 40 or 80 miles bandwidths to state borders. All regressions include controls for tract-level income, electricity prices, electricity prices interacted with income, solar irradiance, population density and population density squared, a battery of tract-level demographic measures, and border-specific polynomials in location relative to border.

Figure A6: Robustness of border discontinuity model to alternative outcomes

(a) Log installations per capita



(b) Log panels per capita

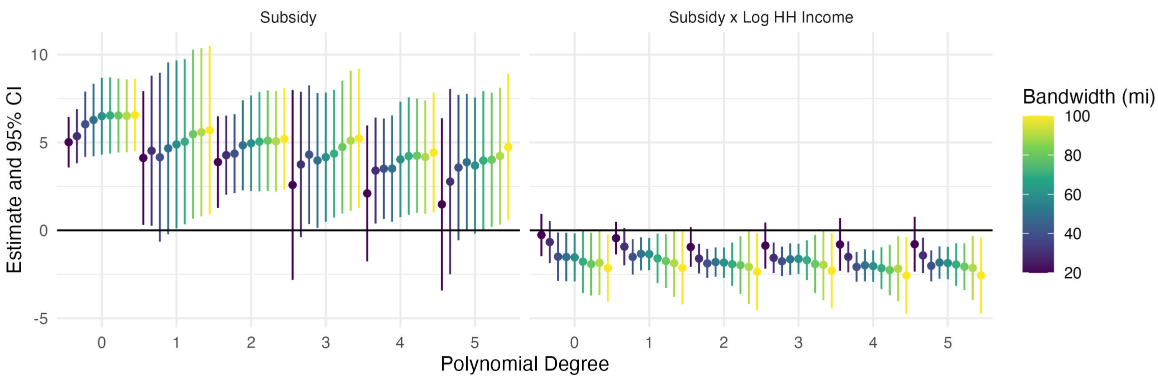
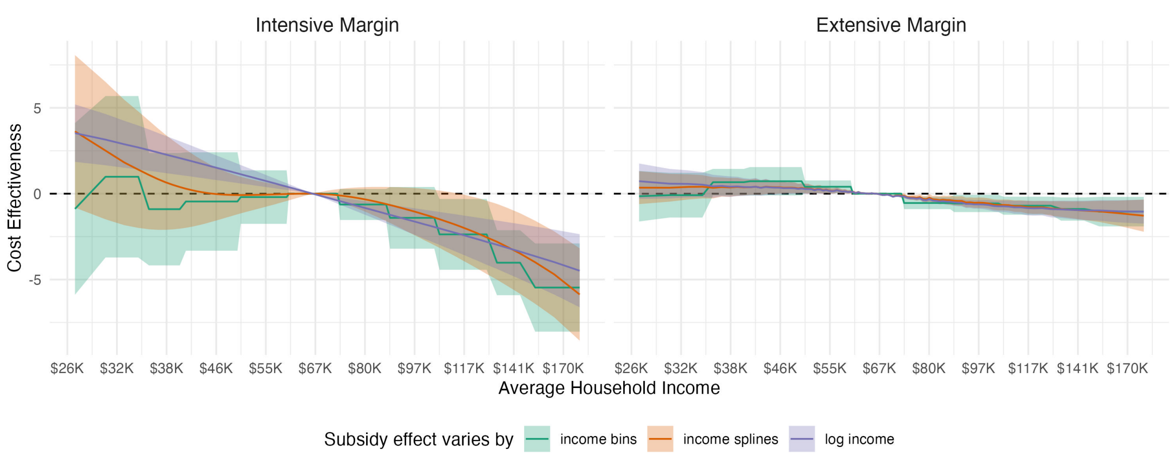
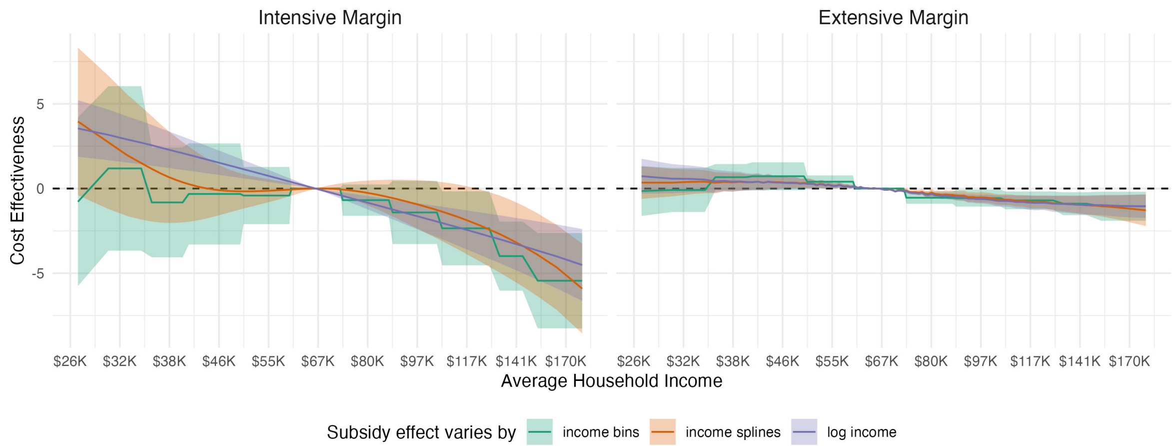


Figure A7: Installations per capita measure of cost effectiveness.



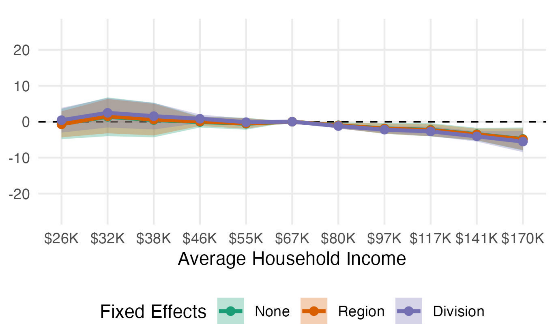
Colors show the results for different manners of allowing the marginal effect of subsidies to vary across income levels, all of which are estimated relative to median income: income bins are 11 bins evenly spaced in log income, income deciles are based on national income distribution, splines are cubic b-splines with 7 knots evenly spaced based on population weighted income, and log income. Demographic, price, and solar controls and division fixed effects included in all regressions. Standard errors are clustered by state.

Figure A8: Panels per capita measure of cost effectiveness.

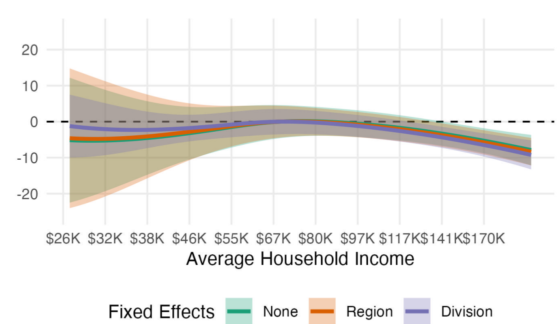


Colors show the results for different manners of allowing the marginal effect of subsidies to vary across income levels, all of which are estimated relative to median income: income bins are 11 bins evenly spaced in log income, income deciles are based on national income distribution, splines are cubic b-splines with 7 knots evenly spaced based on population weighted income, and log income. Demographic, price, and solar controls and division fixed effects included in all regressions. Standard errors are clustered by state.

Figure A9: Installations modelled with a Poisson regression.



(a) Installations as outcome, bins in log income.



(b) Installations as outcome, splines over income.

C Structural Appendix

C.1 Derivation of Partial Elasticity of Installation with Respect to Subsidies

The household value function conditional on installing solar panels is given by the Lagrangian

$$V_i^{m=1} = \max_{c_i, a_i} \sum_{t=1}^{T} \beta^{t-1} \frac{(c_{it})^{1-\gamma}}{1-\gamma} + \phi_i - \sum_{t=1}^{T} \mu_t g_t (a_{it}) - \sum_{t=1}^{T} \lambda_t h_t (c_{it}, a_{it}, a_{it+1})$$

where c_i and a_i are the vectors of household consumption levels and asset levels in all periods, $g_t(a_{it})$ denotes the borrowing inequality constraint in period t, and $h_t(c_{it}, a_{it}, a_{it+1})$ denotes the budget equality constraint in period t.

Let $\bar{V}_i^{m=1} = V_i^{m=1} - \sigma\epsilon_i$ denote household i's value of installing panels less the idiosyncratic preference draw. Note that $\bar{V}_i^{m=1}$ is implicitly a function of subsidies. We can then write the probability of installation as

$$\log P_i = \left(\frac{1}{\sigma}\bar{V}_i^{m=1}\right) - \log\left(\exp\left(\frac{1}{\sigma}\bar{V}_i^{m=1}\right) + \exp\left(\frac{1}{\sigma}\sum_{t=1}^{T}\beta^{t-1}\frac{\left(c_{it}^{m=0}\right)^{1-\gamma}}{1-\gamma}\right)\right).$$

Taking the derivative of $\log P_i$ with respect to s^{Upfront} yields

$$\frac{\partial \log P_i}{\partial s^{\text{Upfront}}} = \frac{1}{\sigma}\frac{\partial \bar{V}_i^{m=1}}{\partial s^{\text{Upfront}}}(1-P_i). \tag{20}$$

Using the notation from above, the first year's borrowing equality constraint is

$$\begin{aligned} &h_t(c_{i1}, a_{i1}, a_{i2}) = c_{i1} + a_{i2} + m_i p_j^{\text{Ins}}(N_i) \\ -&\left(y_i - \tau(y_i) + (1+r)a_{i1} + m_i\left(N_i A_{it} p_j + s_i^{\text{Upfront}}\left(p_j^{\text{Ins}}(N_i), \tau(y_i)\right) + s_i^{\text{Flow}}\left(p_j^{\text{Ins}}(N_i), A_{it}\right)\right)\right). \end{aligned}$$

Therefore, the first-order condition of the Lagrangian with respect to c_1 yields

$$\left(c_{i1}^{m=1}\right)^{-\gamma} = \lambda_1. \tag{21}$$

Further, by the envelope theorem, we know what

$$\frac{\partial \bar{V}_i^{m=1}}{\partial s^{\text{Upfront}}} = \frac{\partial V_i^{m=1}}{\partial s^{\text{Upfront}}} = \lambda_1. \tag{22}$$

Combining equations (20), (21), and (22) yields

$$\frac{\partial \log P}{\partial s^{\text{Upfront}}} = \frac{\left(c_{i1}^{m=1}\right)^{-\gamma}}{\sigma}\left(1 - P_i\right).$$

C.2 Additional Data and Estimation Details

Interest and Discount Rates We follow Sexton et al. (2021) and set a real interest rate of 5% and assume panels have a life of 20 years. We also set $T = 20$. We assume a household discount rate of $\beta = \frac{1}{1+r}$. De Groote and Verboven (2019), who estimate a discount factor by estimating responses of residential solar demand to the introduction of a generous subsidy for future solar production in Belgium. They find that households' implicit real interest rate in evaluating future benefits greatly exceeds the real market interest rate. It would be straightforward to repeat our analysis with alternative values of the household discount rate.

Installation Size We parameterize the number of panels conditional on installation as $N_i = \kappa X_i$, where κ is a vector of parameters to be estimated and X_i is a vector including a constant term, household income, and tract-level college and democrat share. We estimate the vector of parameters κ jointly with the other structural parameters.

Federal Income Taxes For the federal income tax function $\tau(y_i)$, we utilize the functional form used by Heathcote, Storesletten, and Violante (2017), which has been shown to effectively replicate many features of the US income tax code.[41] The functional form for the total tax burden is given by

$$\tau(y_i) = y_i - \lambda y_i^{1-\tau},$$

where τ is a parameter that dictates the progressivity of the tax schedule, and λ is a parameter that dictates the overall level of taxes. We use the values of τ and λ estimated by Guner, Kaygusuz, and Ventura (2014), who estimate these parameters using microdata from the IRS. We use their estimates for all married households.

Subsidies We use data on subsidies Sexton et al. (2021), which are assembled from data from the Database of State Incentives for Renewables & Efficiency (DSIRE). We

[41] See e.g. Guner, Kaygusuz, and Ventura (2014).

assume that upfront subsidies are given by the sum of subsidies from state and federal investment tax credits and sales tax rebates. We assume that flow subsidies are given by the sum of solar renewable energy certificates, other production-based subsidies, and property tax rebates.

Formally, let s^{Fed} denote the portion of the Federal Investment Tax Credit that is refundable in year 1. This refundable portion equals the minimum of the household's tax burden and 30% of the cost of installation. We can write this as

$$s^{\text{Fed}} = \max\left\{0, \min\left\{0.3p_j^{\text{Ins}}\left(N_i\right), \tau\left(y_i\right)\right\}\right\}.$$

Let s_j^{Cost} denote state cost-based subsidies—subsidies which pay a fraction of the cost of installation, let SalesTaxRebate_j be a dummy variable indicating state j offers a sales tax exemption, and let SalesTaxRate_j denote the average sales tax in state j. We can write

$$s_i^{\text{Upfront}} = p_j^{\text{Ins}}\left(N_i\right) \times \left(s_j^{\text{Cost}} + \text{SalesTaxRebate}_j \times \text{SalesTaxRate}_j\right) + s^{\text{Fed}}. \tag{23}$$

Let s_j^{Kwh} denote the production-based subsidies, subsidies which pay per kWh of electricity produced, let PropTaxRebate_j be a dummy variable indicating state j offers a property tax exemption, and let PropTaxRate_j denote the average property tax rate in state j. Flow subsidies are given by

$$s_{it}^{\text{Flow}} = \left(1-\delta\right)^{t-1} p_j^{\text{Ins}}\left(N_i\right) \times \text{PropTaxRebate}_j \times \text{PropTaxRate}_j + s_j^{\text{Kwh}} N_i A_{it}$$

where δ is the depreciation rate of solar panels.

C.3 Construction of Tract-Level Income Distributions

We construct tract-level income distributions for homeowners in two steps. In the first step, we construct tract-level income distributions for all households, including non-homeowners, using tract-level data on average household income, Gini coefficient, and number of households. In the second step, we estimate the the joint distribution between homeownership rates and income using household-level data from the 2015 ACS. We then combine these estimates of homeownership rates with our tract-level income distributions to construct tract-level income distributions for homeowners.

More specifically, in the first step, we assume household income in each tract follows a log-normal distribution and choose the mean and variance of each tract's income distribution to match the tract's average income and Gini coefficient.[42] This allows us to construct the unconditional household income distribution for each tract in our data.

In the second step, we estimate the relationship between household income and the probability that a household is a homeowner separately for each state. Letting Own_i denote that a given household i in the ACS owns their home, we regress

$$\text{Own}_i = f_s(y_i) + \varepsilon_i$$

where y_i is household i's income, and $f_s(y_i)$ is state-specific linear spline in household income with knots at 10, 20, 30, 50, 75, 100, 150, and 200 thousand dollars. We then take each household in the model's predicted value from these regressions to calculate the probability that they are a homeowner. This gives us the joint distribution of income and homeownership.

Finally, we multiply the unconditional distribution of income across tracts unconditional distribution of income across tracts constructed in the first step with this homeownship probability to construct our tract-level income distributions conditional on homeownership.

C.4 Ordered Logit Model of Initial Assets

We discretize the initial asset distribution into $N = 20$ mass points ranging from the minimum asset level, $\bar{a}$, to $\bar{a} + 1,000,000$. Let $n = 1, ..., N$ index mass points, and let $\hat{a}_n$ denote the asset level associated with the nth mass point. We assume the probability that household i has initial assets associated with the nth mass point takes the form of ordered logit probabilities. Therefore, the probabilities that household i is associated

[42] Numerous studies find that the income distribution in the United States is approximately log-normal (see e.g. Battistin, Blundell, and Lewbel (2009).)

		Estimate	Standard Error
Dispersion of Idiosyncratic Utility	σ	2.54	0.26
Curvature of Utility	γ	0.35	0.05
Nonpecuniary Value of Installations			
Constant	ϕ_0	-8.97	0.96
Percent College	ϕ_{Coll}	3.28	0.65
Percent Democrat	ϕ_{Pol}	-2.32	0.57
Size of Installation Parameters			
Constant	κ_0	15.36	1.54
Percent College	κ_{Coll}	-2.32	0.30
Percent Democrat	κ_{Pol}	-5.42	0.55
Demeaned Log Income	κ_{Inc}	0.96	0.15

Table A7: Parameter estimates. Bootstrap standard errors in parenthesis.

with each discrete asset level are given by

$$Pr_i(\hat{a}_N) = \frac{\exp(a_i^\star - \zeta_{N-1})}{1 + \exp(a_i^\star - \zeta_{N-1})}$$
$$Pr_i(\hat{a}_n) = \frac{\exp(a_i^\star - \zeta_{n-1})}{1 + \exp(a_i^\star - \zeta_{n-1})} - \frac{\exp(a_i^\star - \zeta_n)}{1 + \exp(a_i^\star - \zeta_n)} \quad \text{for } n = 2, ...M-1$$
$$Pr_i(\hat{a}_1) = 1 - \frac{\exp(a_i^\star - \zeta_1)}{1 + \exp(a_i^\star - \zeta_1)}$$

where $a_i^\star = \delta y_i$ denotes the expected latent initial assets of household i, δ is a parameter to be estimated, y_i is household income, and ζ's are cut points to be estimated. We constrain that the ζ's to be linear in asset levels

$$\zeta_n = \phi_0 + \phi_1 \hat{a}_n.$$

We estimate the δ and ϕ parameters via indirect inference jointly with the structural parameters.

C.5 Parameter Estimates

Table A7 shows the estimates of the structural parameters with bootstrapped standard errors.

Moment	Data	Simulation
I. Regression of log panels on subsidies and subsidies interacted with log income, controlling for tract characteristics		
Coefficient on subsidies	5.33	5.35
Coefficient on subsidies×log income	-1.92	-1.96
II. Regression of panels per capita on demeaned log income		
Coefficient on demeaned log income	0.25	0.24
Constant	0.21	0.21
III. Average panels per capita across income distribution		
Income quintile 1	0.07	0.09
Income quintile 2	0.13	0.14
Income quintile 3	0.19	0.19
Income quintile 4	0.28	0.26
Income quintile 5	0.40	0.38
Income over 120 thousand	0.41	0.42
IV. Regression of average installation size on characteristics		
Constant	15.73	15.73
Coefficient on college education	-2.21	-2.21
Coefficient on democrat	-5.57	-5.57
Coefficient on demeaned log income	0.64	0.65
V. Regression of panels per capita on percent college education		
Coefficient on college education	0.17	0.47
Constant	0.22	0.21
VI. Regression of panels per capita on percent Democrat		
Coefficient on democrat	0.06	0.01
Constant	0.22	0.22

Table A8: Model fit of targeted moments. The column "Data" gives the value of the moment in the data, while "Simulated" gives the moment calculated in the estimated model.

C.6 Additional Model Fit

Table A8 presents the fit for targeted moments. The model overpredicts the relationship between college share and installations and underpredicts the relationship between average tract income and percent democrat but fits relatively well overall.

C.7 Comparison to Existing Literature

We further assess model validity by simulating natural experiments that design-based papers on solar installation have studied.

Colas and Saulnier (2023) estimate demand for solar panels using a border-discontinuity approach that exploits variation in subsidies on either side of state borders. Their estimates imply that a \$1,000 increase in subsidies for solar panels leads to roughly an 9% increase in solar panel installations. Simulating a \$1,000 increase in subsidies for solar panels in our model leads to a 10.5% increase in installations in our model.

Hughes and Podolefsky (2015) study the effect of the California Solar Initiative rebates by exploiting variation in rebate rates across utility companies. They find that a \$470 increase in total rebate leads to a 10% increase in installations. We replicate this experiment in our model by providing additional upfront subsidies of \$470 to California households. We find this leads to a 4.6% increase in installations in California.

Crago and Chernyakhovskiy (2017) estimate the responsiveness of solar installations to rebates using panel data from the US Northeast and find that increasing rebates by \$1 per watt increases solar panel installations by 47%. We simulate providing the same rebate in these 12 states and find that total installations increase by 44%.

Gillingham and Tsvetanov (2019) estimate the price elasticity of demand for solar panel installations using data from Connecticut and an instrumental variable approach that accounts for excess zeros and unobserved heterogeneity. They find a price elasticity of demand evaluated at the mean installation price equal to -0.65. We simulate the effects of increasing installation prices for households in Connecticut by \$1000 and calculate the implied elasticity. We find an elasticity of demand evaluated at the mean installation price of -0.82.

C.8 Estimation of Marginal Damages of Electricity Production by NERC Region

Let D_{It} denote total environmental damages from electricity production for all plants located within interconnection I in a given hour t. Our estimating equation is given by

$$D_{It} = \sum_{R \in \mathbf{R}_I} \beta_R Load_{Rt} + \alpha_{mh} + \varepsilon_t,$$

where R indexes NERC regions, $\mathbf{R}_I$ is the set of NERC regions in interconnection I, $Load_{Rt}$ is total load in region R in hour t, and α_{mh} are month-by-hour fixed effects. We restrict our sample to 8 AM to 6 PM so that we only measure the marginal damages associated with daytime energy production. The β_R's are the coefficients of interest and measure the marginal damages associated with additional electricity load in region R.

We use data assembled by Holland et al. (2020), who use data on individual power

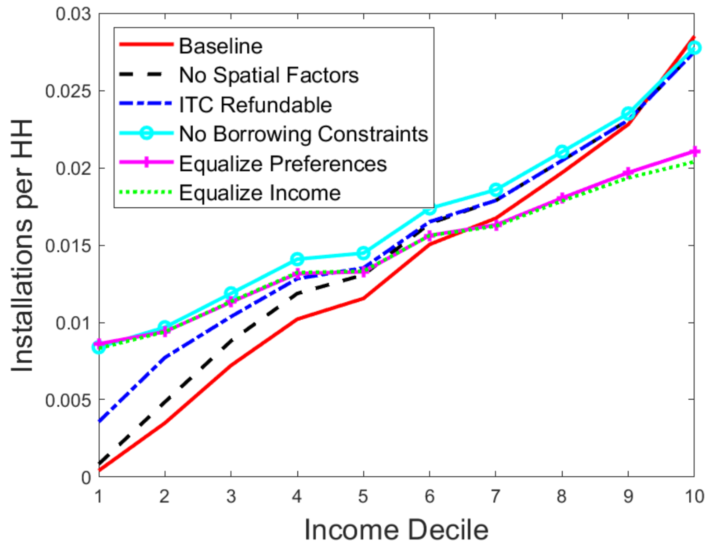


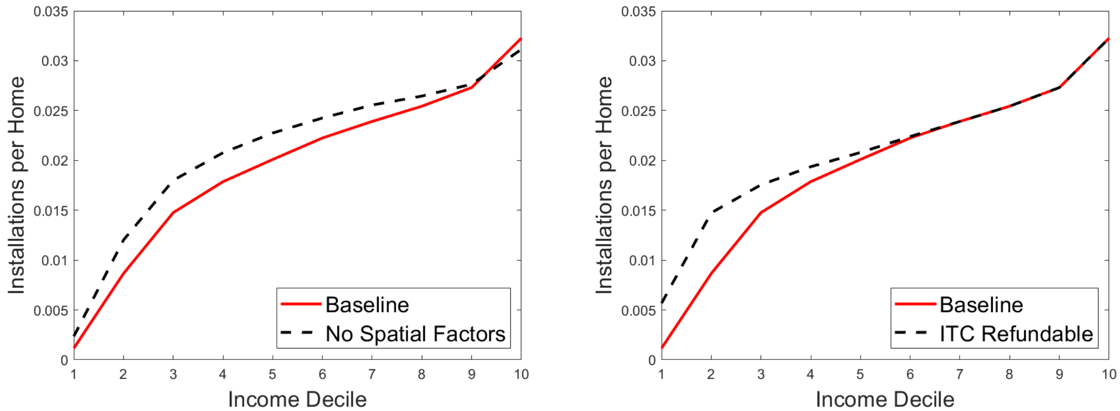
Figure A10: Model-Based Decomposition: All Households. The graph shows the average number of installations per household, including non-homeowners, across deciles of household income over various model specifications. See text for details on each specification.

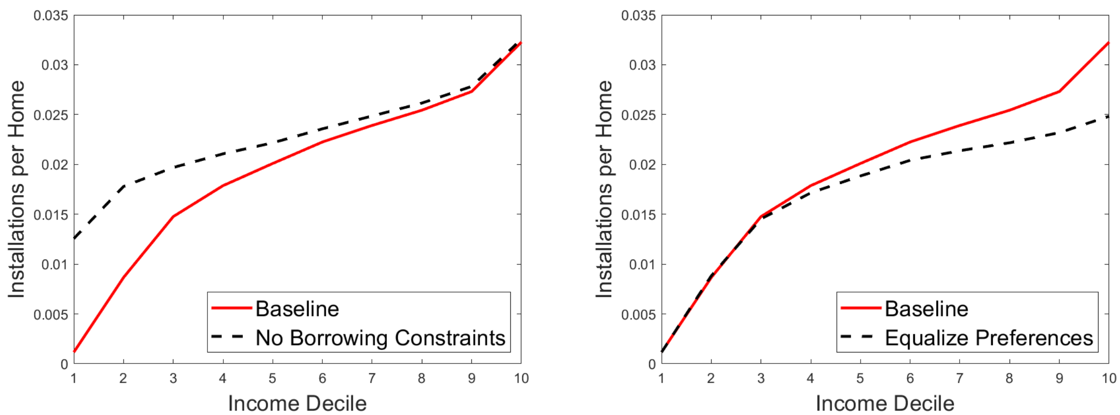
plant production and emissions levels from 2010-2017 from the EPA's Continuous Emissions Monitoring System. Damages are measured as the sum of environmental damages from CO2 and local pollutant emissions. To measure damages associated with CO2 emissions, the authors assume a social cost of carbon valued at \$35.56 per metric ton of CO2 in 2010, which grows at 3 percent annually. To measure environmental damages associated with the emissions of local pollutants, the authors use the AP3 integrated assessment model, which calculates the damages associated with individual pollutants at the plant level.

C.9 Additional Decomposition Results

Figure A10 repeats the decomposition exercise from Section 5.5 but displays installations per household rather than installations per homeowner. Figure A11 repeats the decomposition exercise from Section 5.5, except we do each modification to the model individually rather than cumulatively. Again, the main conclusion is that borrowing constraints and the nonrefundable Federal Investment Tax Credit play important roles in explaining the strong positive relationship between installations and income.

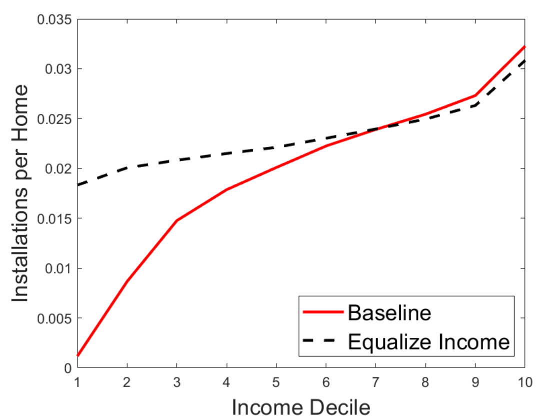
Figure A12 plots the number of *marginal* installations across deciles of household income under various model specifications. The solid red line shows the number of marginal households in the baseline model. The number of marginal households is





(c) Remove Borrowing Constraints

(d) Equality Preferences



(e) Equalize Income

Figure A11: Alternative model decomposition results.

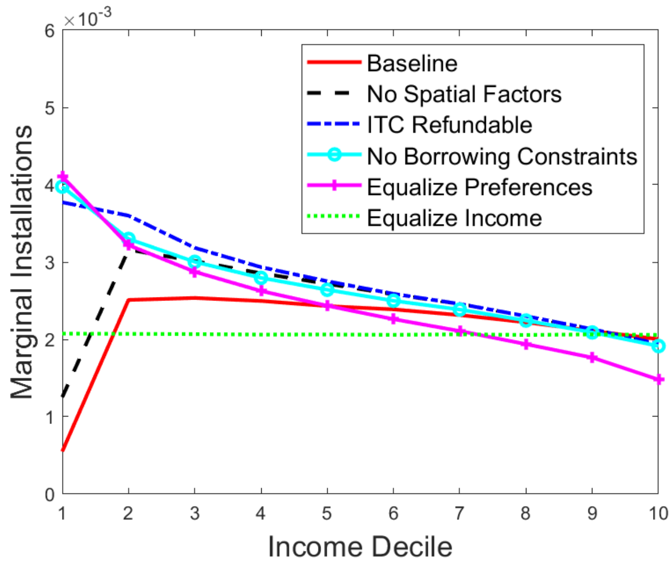


Figure A12: Model-Based Decomposition: Marginal Installations. The graph shows the average number of marginal installations per home across deciles of household income over various model specifications. See text for details on each specification.

initially increasing in income before decreasing. The remaining lines sequentially remove the influence of prices, subsidies, and solar irradiance (black dashed line); make the Federal Investment Tax Credit Refundable (blue dash-dotted line); remove borrowing constraints (cyan dashed line); equalize preferences (solid magenta line); and equalize income (dotted green line).

C.10 Production-Maximizing Optimality Conditions

Let I denote the set of all households, let $P_{i|a}$ denote the probability household i installs panels conditional on having initial assets a, and let $Pr_i(a)$ denote the probability household i has initial asset level a. The government's problem is to maximize solar production

$$\int_{i \in I} \int Pr_i(a)\, P_{i|a} N_i A_i da di$$

subject to the constraint that

$$\int_{i \in I} \int Pr_i(a)\, P_{i|a} s_i da di \leq C,$$

where C is the exogenously-set government budget, and s_i is the present discounted value of subsidies received by household i conditional on installing solar panels. We can write the government's problem as the Lagrangian:

$$\max \int_{i\in I}\int Pr_i\left(a\right)P_{i|a}N_iA_i dadi - \lambda\left(\int_{i\in I}\int Pr_i\left(a\right)P_{i|a}s_i dadi - C\right),$$

where λ is the government's Lagrange multiplier. Let $s^{\text{Inc}}\left(\hat{y}\right)$ denote the income-contingent subsidy for households with income level $\hat{y}$. Taking the first-order condition with respect to $s^{\text{Inc}}\left(\hat{y}\right)$ yields

$$\int_{i\in I(\hat{y})}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}\left(\hat{y}\right)}N_iA_i dadi - \lambda\left(\int_{i\in I(\hat{y})}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}\left(\hat{y}\right)}s_i dadi + \int_{i\in I(\hat{y})}\int Pr_i\left(a\right)P_{i|a}dadi\right)$$

where $I\left(\hat{y}\right)$ is the set of households with income level $\hat{y}$.

Let $M\left(\hat{y}\right) = \int_{i\in I(\hat{y})}\int Pr_i\left(a\right)P_{i|a}dadi$ denote the total number installations by households with income level $\hat{y}$. We can then rewrite the government's optimality condition as

$$\frac{\partial M\left(\hat{y}\right)}{\partial s^{\text{Inc}}\left(\hat{y}\right)} \times \left(\bar{s}\left(\hat{y}\right) - \frac{1}{\lambda}\overline{NA}\left(\hat{y}\right)\right) + M\left(\hat{y}\right) = 0.$$

The term $\frac{\partial M(\hat{y})}{\partial s^{\text{Inc}}(\hat{y})}$ gives the derivative of installations of households with income level $(\hat{y})$. These marginal installations increase government costs, as these households now receive subsidies. This marginal cost is captured by $\bar{s}\left(\hat{y}\right)$, which gives the average subsidy received across all marginal households. This is formally given by

$$\bar{s}\left(\hat{y}\right) = \frac{\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}(\hat{y})}s_i dadi}{\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}(\hat{y})}dadi}.$$

These marginal households are also associated with additional solar production. This is captured by the term $\overline{NA}\left(\hat{y}\right)$, which gives the average solar output per installation for these marginal households. This term is formally given by

$$\overline{NA}\left(\hat{y}\right) = \frac{\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}(\hat{y})}N_iA_i dadi}{\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}(\hat{y})}dadi}.$$

Finally, households at income level $\hat{y}$ who already choose to install panels given the current levels of subsidies would receive additional subsidies if the government increased

$s^{\text{Inc}}(\hat{y})$. The additional government cost associated with more generous subsidies for these inframarginal households is captured by the term $M(\hat{y})$.

C.11 Environmental-Benefit-Maximizing Subsidies

Optimality Conditions Let B_i denote the environmental benefits associated with one kWh of solar electricity produced by household i. Let I denote the set of all households, let $P_{i|a}$ denote the probability household i installs panels conditional on having initial assets a, and let $Pr_i(a)$ denote the probability household i has initial asset level a. The government's problem is to maximize environmental benefits,

$$\int_{i\in I}\int Pr_i(a)\,P_{i|a}N_iA_iB_i\,dadi$$

subject to the constraint that

$$\int_{i\in I}\int Pr_i(a)\,P_{i|a}s_i\,dadi \leq C,$$

where C is the exogenously set government budget, and s_i is the present discount value of subsidies received by household i conditional on installing solar panels.

Similar to Appendix C.10, we can then write government's optimality condition as

$$\frac{\partial M(\hat{y})}{\partial s^{\text{Inc}}(\hat{y})} \times \left(\bar{s}(\hat{y}) - \frac{1}{\lambda}\overline{NAB}(\hat{y})\right) + M(\hat{y}) = 0.$$

The term $\frac{\partial M(\hat{y})}{\partial s^{\text{Inc}}(\hat{y})}$ gives the derivative of installations for households with income level $(\hat{y})$. The marginal cost associated with providing subsidies to these households is captured by $\bar{s}(\hat{y})$, which gives the average subsidy received across all marginal households and is formally given by

$$\bar{s}(\hat{y}) = \frac{\int_{i\in I}\int Pr_i(a)\,\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}(\hat{y})}s_i\,dadi}{\int_{i\in I}\int Pr_i(a)\,\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}(\hat{y})}\,dadi}.$$

These marginal households are also associated with additional environmental benefits. This is captured by the term $\overline{NAB}(\hat{y})$, which gives the average environmental benefits per installation for these marginal households. This term is formally given by

$$\overline{NAB}\left(\hat{y}\right)=\frac{\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}\left(\hat{y}\right)}N_iA_iB_idadi}{\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}\left(\hat{y}\right)}dadi}.$$

Finally, the additional government cost associated with more generous subsidies for these inframarginal households is captured by the term $M\left(\hat{y}\right)$.

Results To measure B_i, the marginal benefits associated with one kWh of solar electricity produced by household i, we estimate NERC-region level marginal damages of electricity production using data from Holland et al. (2020). We describe this estimation procedure in Appendix C.8.

Table A9 shows the results. Each column presents subsidy levels, installation rates, total solar production, and total environmental benefits under a given subsidy scheme. The first column presents these statistics under the current subsidy scheme, the second column presents the production-maximizing subsidy scheme, and the third column presents the environmental-benefits-maximizing subsidy scheme. The optimal subsidy schemes, distribution of installations, production levels, and environmental benefits are very similar for the production-maximizing and environmental-benefit-maximizing subsidy schemes. The environmental-benefit-maximizing subsidy scheme leads to a 3.9% increase in environmental benefits of residential solar relative to the current subsidy scheme.

C.12 Welfare-Maximizing Subsidies

Let B_i denote the environmental benefits associated with one kWh of solar electricity produced by household i and let $V_{i|a}$ denote the lifetime utility of household i conditional on having initial assets a. The government's problem is to maximize

$$\underbrace{\int_{i\in I}\int Pr_i\left(a\right)P_{i|a}V_{i|a}dadi}_{\text{Utilitarian Welfare}}$$

subject to the constraint that

$$\underbrace{\int_{i\in I}\int Pr_i\left(a\right)P_{i|a}s_idadi}_{\text{Fiscal Cost}}-\underbrace{\int_{i\in I}\int Pr_i\left(a\right)P_{i|a}N_iA_iB_idadi}_{\text{Environmental Benefit}}\leq C,$$

	(1) Baseline	(2) Prod Max	(3) Benefit Max
I. Production per HH			
Income Q1	13.2	33.4	32.0
Income Q2	48.5	64.3	63.9
Income Q3	80.2	80.4	81.6
Income Q4	123.5	96.7	97.0
Overall	68.2	69.8	69.8
II. Subsidy Generosity ($1000s)			
Income Q1	8.1	12.0	11.9
Income Q2	9.3	11.4	11.4
Income Q3	9.8	9.8	10.0
Income Q4	10.4	6.9	7.0
III. Relative Production	100.0	102.4	102.3
IV. Relative Benefits	100.0	102.3	102.3

Table A9: Environmental-Benefit-Maximizing Subsidies. Panel I shows the average yearly solar capacity in kWh per household in each income quartile. Panel II shows the average subsidy a household from each income quartile would receive for a solar installation. Panel III shows the total solar production. We scale total production under the baseline simulation to 100. Panel IV shows the total environmental benefits of solar panels. We scale the environmental benefits under the baseline simulation to 100.

where C is the exogenously set maximum net cost.

We can write the constrained welfare maximizing problem as the Lagrangian

$$\max W = \underbrace{\int_{i\in I}\int Pr_i\left(a\right)V_{i|a}dadi}_{\text{Utilitarian Welfare}} - \lambda\left(\underbrace{\int_{i\in I}\int Pr_i\left(a\right)P_{i|a}s_idadi}_{\text{Fiscal Cost}} - \underbrace{\int_{i\in I}\int Pr_i\left(a\right)P_{i|a}N_iA_iB_idadi}_{\text{Environmental Benefit}} - C\right), \quad (24)$$

where λ is the government's Lagrange multiplier.

The system of optimal subsidies must satisfy the government's first-order conditions, which implies

$$\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial V_{i|a}}{\partial s^{\text{Inc}}\left(\hat{y}\right)}dadi - \lambda\left(\int_{i\in I}\int Pr_i\left(a\right)\left(\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}\left(\hat{y}\right)}\left(s_i - N_iA_iB_i\right) + P_{i|a}\right)dadi\right) = 0. \quad (25)$$

Again let $M\left(\hat{y}\right) = \int_{i\in I(\hat{y})}\int Pr_i\left(a\right)P_{i|a}dadi$ denote the total number installations by households with income level $\hat{y}$. We can then rewrite the government's optimality condition as

$$\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial V_{i|a}}{\partial s^{\text{Inc}}\left(\hat{y}\right)}dadi + \lambda\left(\frac{\partial M\left(\hat{y}\right)}{\partial s^{\text{Inc}}\left(\hat{y}\right)} \times \left(\overline{NAB}\left(\hat{y}\right) - \bar{s}\left(\hat{y}\right)\right) - M\left(\hat{y}\right)\right) = 0.$$

The term $\frac{\partial M(\hat{y})}{\partial s^{\text{Inc}}(\hat{y})}$ gives the derivative of installations of households with income level $(\hat{y})$. These marginal installations lead to environmental benefits but also increase government cost. This government cost is captured by $\bar{s}\left(\hat{y}\right)$, which gives the average subsidy received across all marginal households. This is formally given by

$$\bar{s}\left(\hat{y}\right) = \frac{\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}(\hat{y})}s_idadi}{\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}(\hat{y})}dadi}.$$

The increase in environmental benefits is given by $\overline{NAB}\left(\hat{y}\right)$, which gives the average environmental benefits of marginal installations for income level $\hat{y}$. This is formally given by

$$\overline{NAB}\left(\hat{y}\right)=\frac{\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s^{\text{Inc}}(\hat{y})}N_iA_iB_idadi}{\int_{i\in I}\int Pr_i\left(a\right)\frac{\partial P_{i|a}}{\partial s_j^{\text{Inc}}(\hat{y})}dadi}.$$

Using the envelope theorem, we know that

$$\frac{\partial V_{i|a}}{\partial s_j^{\text{Inc}}\left(\hat{y}\right)}=P_{i|a}\frac{\partial u_i}{\partial c_{i1}}\left(c_{i1|a}^{m=1}\right)$$

where $\frac{\partial u_i}{\partial c_{i1}}\left(c_{i1|a}^{m=1}\right)$ gives the marginal utility of consumption with respect to consumption in year 1, evaluated at the optimal consumption level in year 1 conditional on installation and having initial asset level a.

We can then write the government's optimality condition in a similar form to that in Colas, Findeisen, and Sachs (2021) as

$$\frac{\partial M\left(\hat{y}\right)}{\partial s^{\text{Inc}}\left(\hat{y}\right)}\times\left(\overline{NAB}\left(\hat{y}\right)-\bar{s}\left(\hat{y}\right)\right)-M\left(\hat{y}\right)\left(1-\bar{W}\left(\hat{y}\right)\right)=0,$$

where $\bar{W}\left(\hat{y}\right)$ gives the money-metric average social welfare weights of households of income group $\hat{y}$ who install solar panels. This is formally given by

$$\bar{W}\left(\hat{y}\right)=\frac{1}{\lambda}\frac{\int_{i\in I}\int Pr_i\left(a\right)P_{i|a}\frac{\partial u_i}{\partial c_{i1}}\left(c_{i1|a}^{m=1}\right)dadi}{\int_{i\in I}\int Pr_i\left(a\right)P_{i|a}dadi}.$$

Intuitively, this gives the average social welfare increase associated with an additional unit of consumption for households who already choose to install solar panels.

C.13 Stochastic Income Details

C.13.1 Model Details

Timing Households begin the model with an initial persistent component of income z_{i0}. At the beginning of the model, they receive their idiosyncratic preferences for panels, ϵ_i, and make a once and for all decision of whether to install solar panels ($m_i \in \{0,1\}$).

After making this decision, households receive the persistent and transitory earnings shocks, η_{it}, and ε_{it}. Households then make a consumption and savings decision, taking an expectation over future earnings shocks. Each year, households continue to receive that year's two shocks and make consumption and earnings decisions.

Model Solution Let $V_{it}(\Omega_{it}, \eta_{it}, \varepsilon_{it})$ denote household i's value function in period t, conditional on state space Ω_{it}, and earnings η_{it}, and ε_{it}. The state space consists of lagged value of the persistent component, z_{it-1}, assets a_{it}, the amount of carried over federal tax credits s_i^{Carry}, value of electricity produced $\left(m_i N_i A_i \left(p_j + s_j^{\text{kwh}}\right)\right)$, and amount of property tax rebate. Note that this value function can be used for both households who choose to install solar panels and for those who do not by setting the appropriate state space variables to 0.

After making the initial installation decision, the household's value function is given by

$$V_{it}(\Omega_{it}, \eta_{it}, \varepsilon_{it}) = \max_{c_{it}} \frac{(c_{it})^{1-\gamma}}{1-\gamma} + \beta \mathbb{E}\left[V_{it+1}(\Omega_{it+1}, \eta_{it+1}, \varepsilon_{it+1} | c_{it})\right]$$

subject to the budget constraint ((7) and (8)) and the borrowing constraint (9). The expectation is taken over η_{it+1} and ε_{it+1}.

Let $\Omega_{i1}^{m=0}$ denote the individual's state space in year 1 conditional on not installing solar panels, and let $\Omega_{i1}^{m=1}$ denote the state space conditional on installing solar panels. The household chooses to install solar panels if

$$\mathbb{E}\left[V_{it}\left(\Omega_{i1}^{m=1}, \eta_{i1}, \varepsilon_{i1}\right)\right] + \phi_i \geq \mathbb{E}\left[V_{it}\left(\Omega_{i1}^{m=0}, \eta_{i1}, \varepsilon_{i1}\right)\right].$$

Given that ϵ_i has a logit distribution, the probability household i installs solar panels is given by

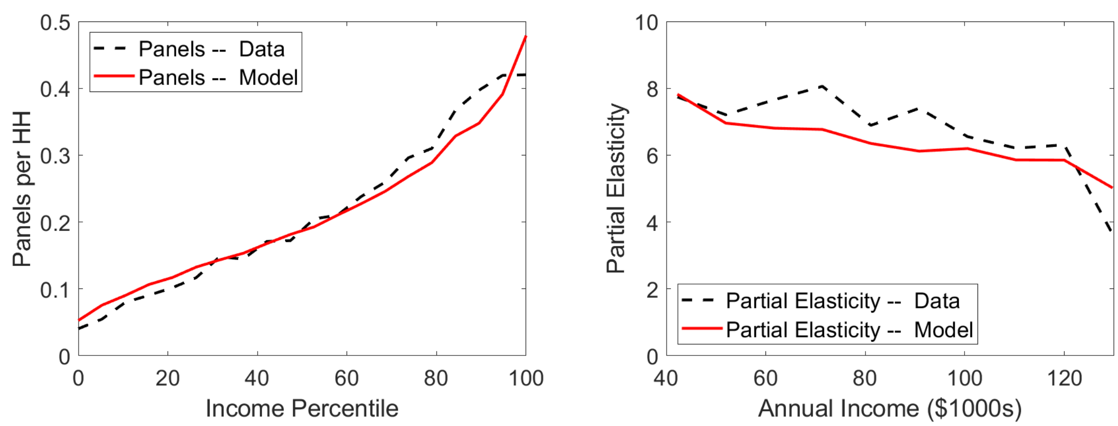
$$P_i = \frac{\exp\left(\frac{1}{\sigma}\mathbb{E}\left[V_{it}\left(\Omega_{i1}^{m=1}, \eta_{i1}, \varepsilon_{i1}\right)\right] + \bar{\phi}_i\right)}{\exp\left(\frac{1}{\sigma}\mathbb{E}\left[V_{it}\left(\Omega_{i1}^{m=1}, \eta_{i1}, \varepsilon_{i1}\right)\right] + \bar{\phi}_i\right) + \exp\left(\frac{1}{\sigma}\mathbb{E}\left[V_{it}\left(\Omega_{i1}^{m=0}, \eta_{i1}, \varepsilon_{i1}\right)\right]\right)}.$$

Calibration We set $\sigma_\eta = 0.022$, $\sigma_\varepsilon = 0.057$, and $\rho = 0.984$, based on the estimates from Panel B of Table 1 from Storesletten, Telmer, and Yaron (2004). We assume the initial value of the persistent component of earnings, z_{i0}, follows a normal distribution with tract-specific means and variances such that income in year 1 also follows a log-normal distribution. We choose the tract-specific means and variances of z_{i0} such that the income distribution in year $t = 1$ matches tract-level average income and Gini coefficients.

We assume that the income-contingent subsidies are awarded based on a household's income in year $t = 0$. We assume there is no transitory income shock in year 0 such

		Estimate
Dispersion of Idiosyncratic Utility	σ	2.51
Curvature of Utility	γ	0.35
Nonpecuniary Value of Installations		
Constant	ϕ_0	-8.56
Percent College	ϕ_{Coll}	2.84
Percent Democrat	ϕ_{Pol}	-2.05
Size of Installation Parameters		
Constant	κ_0	15.32
Percent College	κ_{Coll}	-2.33
Percent Democrat	κ_{Pol}	-5.52
Demeaned Log Income	κ_{Inc}	1.01

Table A10: Parameter estimates for model with stochastic income



(a) Income (b) Partial Elasticity

Figure A13: Model Fit: Stochastic Income. Panel (a) shows panels per household across percentiles of tract-level income in the model and data. The red line shows the panels per household in the data, while the black dashed line shows the simulated panels from the estimated model. These are estimated using the same set of controls as Figure 4.

that $\log y_{i0} = z_{i0}$.

C.13.2 Results

Table A10 presents the parameter estimates for the model with stochastic income. The parameter estimates are similar to the baseline estimates.

Figure A13 shows model fit. The first panel shows panels per household across percentiles of tract-level income in the model and data. The red line shows the panels per household in the data while the black dashed line shows the simulated panels from the estimated model. The second panel shows partial elasticities across income levels in the model and data. These are estimated using the same set of controls as Figure 4.

	(1) Baseline	(2) Prod Max	(3) Utility Max
I. Production per HH			
Income Q1	5.1	19.7	28.9
Income Q2	52.6	65.6	69.7
Income Q3	81.4	84.8	81.6
Income Q4	125.4	102.4	93.1
Overall	68.1	69.6	69.5
II. Subsidy Generosity ($1000s)			
Income Q1	8.0	11.7	13.3
Income Q2	9.3	11.1	11.6
Income Q3	9.8	10.2	9.8
Income Q4	10.4	7.7	6.5
III. Relative Production	100.0	102.1	102.0

Table A11: Results: Stochastic Income. Panel I shows the average yearly solar capacity in kWh per household in each income quartile. Panel II shows the average subsidy a household from each income quartile would receive for a solar installation. Panel III shows the total solar production. We scale total production under the baseline simulation to 100.

Table A11 has the main results with stochastic income. The optimal subsidy schemes, distribution of installations, and production levels are similar to those under the baseline model.