NBER WORKING PAPER SERIES

A QUANTITATIVE ANALYSIS OF SUBSIDY COMPETITION IN THE U.S.

Ralph Ossa

Working Paper 20975
http://www.nber.org/papers/w20975

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2015

I am grateful to Kerem Cosar, Jonathan Dingel, Owen Zidar, and seminar participants at various universities and conferences for very helpful comments and suggestions. The usual disclaimer applies. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

A Quantitative Analysis of Subsidy Competition in the U.S.
Ralph Ossa
NBER Working Paper No. 20975
February 2015, Revised October 2017
JEL No. F12,F13,R12,R58

ABSTRACT

I use a quantitative economic geography model to explore subsidy competition among U.S. states. I ask what motivates state governments to subsidize firm relocations and quantify how strong their incentives are. I also characterize fully non-cooperative and cooperative subsidy choices and assess how far away we are from these extremes. I find that states have strong incentives to subsidize firm relocations in order to gain at the expense of other states. I also find that observed subsidies are closer to cooperative than non-cooperative subsidies but the potential losses from an escalation of subsidy competition are large.

Ralph Ossa
University of Zurich
Department of Economics
Schoenberggasse 1
8001 Zuerich
Switzerland
and NBER
ralph.ossa@econ.uzh.ch

1 Introduction

U.S. state and local governments spend substantial resources on subsidies competing for mobile firms. According to a database from the W.E. Upjohn Institute for Employment Research, the costs of such subsidies have more than tripled since 1990 reaching a total of $45 billion in 2015. This figure is equivalent to around 30 percent of state and local business tax revenue and adds up all subsidies that are commonly available to medium and medium-large firms. They include property tax abatements, customized job training subsidies, investment tax credits, and research and development tax credits, among other things.[1]

In this paper, I provide a first comprehensive quantitative analysis of this subsidy competition in the U.S.. I first ask what motivates governments to subsidize firm relocations and quantify how strong their incentives are. I then characterize fully non-cooperative and cooperative subsidy choices and assess how far away we are from these extremes. By doing so, I aim to make sense of a widely used policy intervention and inform the surrounding policy debate.

I pursue this analysis in the context of a quantitative economic geography model which I calibrate to U.S. states. Influenced by the trade policy literature, I calculate optimal subsidies, Nash subsidies, and cooperative subsidies and then compare them to observed subsidies. Optimal subsidies are the subsidies states would offer if they did not have to fear any retaliation and shed light on the incentives states have. Nash subsidies and cooperative subsidies then characterize the fully non-cooperative and cooperative subsidy choices thereby capturing the worst-case and best-case scenarios.

I find that states have strong incentives to subsidize firm relocations in order to gain at the expense of other states. Optimal subsidies average $14.9 billion, would raise real income by an average 2.2 percent in the subsidy imposing state, and would lower real income by an average -0.2 percent in all other states. I also find that observed subsidies are much closer to cooperative than non-cooperative subsidies but that the potential costs of an escalation of subsidy competition are large. In particular, moving from observed subsidies to Nash

[1] The database is called Panel Database on Business Incentives and it is documented in Bartik (2017). Earlier estimates put the annual subsidy costs at $46.8 billion in 2005 (Thomas, 2011) and $80.4 billion in 2012 (Story et al., 2012).

subsidies would cost on average -1.1 percent of real income while moving to cooperative (i.e. zero) subsidies would only improve welfare minimally.

The key mechanism in my analysis is an agglomeration externality in the New Economic Geography tradition which derives from an interaction of internal increasing returns and trade costs. In particular, consumers benefit from being close to firms because this gives them access to cheaper final goods. Similarly, firms benefit from being close to firms because this gives them access to cheaper intermediate goods. By subsidizing firm relocations, states try to foster local agglomeration at the expense of other states so that their subsidies are beggar-thy-neighbor policies.

In my quantification, I try to strike a balance between transparency and realism to be able to clearly illustrate the main mechanisms and yet obtain broadly credible quantitative results. Analytical results are notoriously hard to derive in economic geography models so that the quantitative analysis is also meant to convey more fundamental conceptual points. While this means that I have to make compromises, it strikes me as a natural approach in this case, which certainly goes well beyond the arbitrary numerical examples that have long dominated the economic geography literature.

I am not aware of any comparable analysis of non-cooperative and cooperative policy equilibria in a spatial environment. Most closely related is, perhaps, the recent work by Fajgelbaum et al (2016) who use a quantitative economic geography model to study state taxes as a source of spatial misallocation in the United States. However, they only consider the implications of exogenous changes in state taxes and do not attempt to solve for non-cooperative or cooperative policy equilibria. The same basic point applies to well-known earlier contributions to the quantitative place-based policy literature such as Gaubert (2014) and Suarez Serrato and Zidar (2016).[2]

The optimal subsidy argument I develop in the paper builds on the insight of Venables (1987) that governments have an incentive to exploit the agglomeration economies backward and forward linkages bring about. I have already explored the implications of it for tariff wars

[2] Gaubert (2014) quantifies the aggregate effects of subsidies given by the national government to lagging regions in France. Suarez Serrato and Zidar (2016) estimate the incidence of state corporate taxes on the welfare of workers, landowners, and firms in the U.S.. See also Greenstone et al (2010) and Kline and Moretti (2014) for related empirical analyses.

in a series of earlier papers (Ossa, 2011; Ossa, 2012; Ossa, 2014) and also draw on some of the methods I developed there. Having said this, there are some fundamental differences between tariff wars and subsidy wars. The most striking one is that subsidy wars can potentially improve overall welfare because the local spillovers which make subsidy wars tempting also bring about allocative inefficiencies which subsidies can correct.[3,4]

The remainder of the paper is organized as follows. In section 2, I lay out the theoretical framework describing the basic setup, the equilibrium for given subsidies, the general equilibrium effects of subsidy changes, and the agglomeration and dispersion forces at work. In section 3, I turn to the calibration, explaining how I choose the model parameters, what adjustments I make to the model, and how I deal with possible multiplicity. In section 4, I perform the main analysis, exploring the welfare effects of subsidies, optimal subsidies, Nash subsidies, and cooperative subsidies.

2 Framework

The theoretical framework is in the New Economic Geography tradition of Krugman (1991) and Krugman and Venables (1995). It emphasizes agglomeration economies resulting from forward and backward linkages which arise endogenously from the interaction of firm-level increasing returns, transport costs, and factor mobility. The main intuition is that workers want to be close to firms and firms want to be close to firms in order to have cheaper access to goods for final and intermediate use. These agglomeration economies have a beggar-thy-neighbor character which is what governments then exploit.

This formulation of agglomeration economies has a number of attractive features, as discussed extensively in the related literature. For example, Fujita et al (2001) emphasize that it does not simply assume agglomeration economies with reference to imprecise notions such as

[3] As I discuss in detail later on, the abovementioned -1.1 percent real income losses associated with an escalation of subsidy competition are calculated relative to a benchmark in which all allocative inefficiencies are eliminated by the federal government. Absent this intervention, a subsidy war would actually increase real incomes in all states.

[4] My analysis is also related to the tax competition literature following Oates (1972) which emphasizes fiscal externalities. The main difference is that in my model governments do not want to attract firms for fiscal reasons but because they generate local spillover effects. Having said this, Baldwin et al (2005) analyze tax competition in a range of stylized New Economic Geography models which also feature some of the mechanism I emphasize.

localized spillover effects but actually derives them as an endogenous model outcome. Also, empirical studies such as Handbury and Weinstein (2015) provide direct evidence supporting its underlying mechanism by showing that larger regions tend to have lower variety-adjusted price indices.

Having said this, this New Economic Geography model has an isomorphic external increasing returns representation as one might suspect from the work of Allen and Arkolakis (2014). In particular, it can also be interpreted as a perfectly competitive Armington (1969) model with factor mobility in which local productivity is simply assumed to be increasing in local economic activity. In that sense, it can really capture all of the famous Marshallian agglomeration forces deriving from specialized inputs, thick labor markets, and technological spillovers.

2.1 Basic setup

The country is populated by workers who can freely move across regions. They consume final goods and residential land and have location preferences which have an idiosyncratic component. Goods are produced by an endogenous number of monopolistically competitive firms from labor, capital, commercial land, and intermediate goods. Capital is freely mobile across regions, land can be freely put to residential or commercial use, and input-output linkages are of the roundabout form. The total supply of labor and capital is fixed at the national level and the total supply of land is fixed at the regional level.

2.1.1 Preferences

Concretely, the utility of worker v living in region j is given by:

$$\begin{aligned} U_{jv} &= U_j u_{jv} \qquad (1)\\ U_j &= \frac{A_j}{L_j}\left(\frac{T_j^R}{\mu}\right)^{\mu}\left(\frac{C_j^F}{1-\mu}\right)^{1-\mu} \\ C_j^F &= \left(\sum_{i=1}^{R}\int_0^{M_i} c_{ij}^F\left(\omega_i\right)^{\frac{\varepsilon-1}{\varepsilon}}d\omega_i\right)^{\frac{\varepsilon}{\varepsilon-1}} \\ u_{jv} &\sim Frechet\left(1,\sigma\right) \end{aligned}$$

where U_j is its common and u_{jv} is its idiosyncratic component. U_j aggregates amenities A_j, residential land T_j^R, and final goods consumption C_j^F in a Cobb-Couglas fashion with a land-expenditure-share μ. The formula is divided by the local number of workers L_j to express everything in per-capita terms. C_j^F is a CES aggregate of M_i differentiated varieties from each of the R regions with an elasticity of substitution $\varepsilon > 1$. u_{jv} is drawn from a Frechet distribution in an iid fashion and σ is an inverse measure of the dispersion of workers' idiosyncratic location preferences.

While I include land purely for quantitative realism, the idiosyncratic location preferences play a more central role. In particular, they ensure that the common component of utility does not necessarily equalize across space thereby introducing a meaningful sense in which regions can benefit at the expense of other regions. Together, these two ingredients also give rise to the two main congestion forces in the model, namely rising land prices and deteriorating worker-region preference mismatch. As we will see, this mismatch also has interesting implications for the welfare effects of interregional transfer payments.

2.1.2 Technology

Varieties are uniquely associated with firms and produced with the following technology:

$$\begin{aligned} q_j &= \varphi_j \left(z_j - f_j\right) && (2) \\ z_j &= \frac{1}{M_j} \left(\frac{1}{\eta} \left(\frac{L_j}{\theta^L}\right)^{\theta^L} \left(\frac{K_j}{\theta^K}\right)^{\theta^K} \left(\frac{T_j^C}{\theta^T}\right)^{\theta^T} \right)^{\eta} \left(\frac{C_j^I}{1-\eta}\right)^{1-\eta} \\ C_j^I &= \left(\sum_i \int_0^{M_i} c_{ij}^I \left(\omega_i\right)^{\frac{\varepsilon-1}{\varepsilon}} d\omega_i \right)^{\frac{\varepsilon}{\varepsilon-1}} \end{aligned}$$

where z_j is an aggregate input which gets turned into output q_j with productivity φ_j after subtracting fixed costs f_j. z_j combines labor L_j, capital K_j, commercial land T_j^C, and intermediate goods C_j^I in a nested Cobb-Douglas fashion with η being the share of value added in gross production and θ_s^L, θ_s^K, and θ_s^T, $\theta_s^L + \theta_s^K + \theta_s^T = 1$, the shares of value added accruing to labor, capital, and land, respectively. The formula gets divided by the number of firms M_j to express everything in per-firm terms. C_j^I is the same CES aggregate over individual varieties as C_j^F above.

Having multiple factors with varying amounts of effective mobility is important for my results.[5] As I will describe in more detail shortly, local governments provide subsidies to local firms which they finance through local labor taxes. For such subsidies to affect the location of economic activity, it is important that there is a more mobile factor than the one that gets taxed. As is easy to show, they would do nothing but raise the before-tax wage by the amount of the tax/subsidy if labor was the only factor of production, thereby leaving incentives completely unchanged.

2.1.3 Government

I distinguish between a non-cooperative and a cooperative policy regime. In the non-cooperative regime, local governments choose local subsidies to maximize local expected utility, which can be written as $E\left(U_{jv} \middle| \text{ living in } j\right)$. In the cooperative regime, the federal government chooses all subsidies to maximize national expected utility, which is given by $E\left(\max_j \{U_{jv}\}\right)$. National expected utility is defined as the expected value of the maximum of all local utilities since workers are freely mobile across regions and choose whichever one offers them the highest utility.

Since subsidy changes induce workers to re-optimize their location choices, local expected utility can in principle be defined over the set of ex-ante or ex-post local residents. I adopt the ex-ante definition in most of what follows because it strikes me as the more natural one. The most obvious reason is that local policy changes get voted on by current and not future residents of the location. Moreover, we will see that this assumption implies that local governments act (almost) as if they maximized local employment which resonates nicely with the rhetoric of real world policy debates.

While I am therefore quite comfortable with this assumption, I also want to be clear that it is not an innocuous one. In particular, it is easy to verify that the local expected utility of ex-post local residents is actually equalized across locations and equal to the national expected utility. This implies that local governments would simply maximize national welfare if they maximized the expected utility of ex-post local residents in which case there would no longer

[5] While labor is freely mobile across regions, the idiosyncratic location preferences act like a mobility cost. Hence, capital is effectively the most mobile factor in this environment, followed by labor and then land.

be any meaningful difference between the non-cooperative regime and the cooperative regime.

Formally, maximizing the local expected utility of ex-ante local residents is equivalent to maximizing the common component of local utility, U_j. Using the properties of the Frechet distribution, it is easy to show that maximizing the expected utility of national residents is equivalent to maximizing $\left(\sum_{i=1}^{R} U_i^{\sigma}\right)^{\frac{1}{\sigma}}$. With that in mind, I will refer to changes in U_j as changes in local welfare and changes in $\left(\sum_{i=1}^{R} U_i^{\sigma}\right)^{\frac{1}{\sigma}}$ as changes in national welfare in the following. For future reference, I summarize the objective functions of the local and federal governments as:

$$\begin{aligned} G_j^{loc} &= U_j \\ G^{fed} &= \left(\sum_{i=1}^{R} U_i^{\sigma}\right)^{\frac{1}{\sigma}} \end{aligned} \tag{3}$$

To preempt any confusion, let me reiterate that U_j is just amenity adjusted per-capita consumption. As we will see shortly, this then implies that U_j also corresponds to amenity adjusted per-capita real income. For given amenities, local welfare changes can therefore also be interpreted as local per-capita consumption or real income changes. As a result, I use the expressions changes in local welfare, changes in local per-capital consumption, and changes in local per-capita real income interchangeably in the following when discussing the local welfare effects of subsidies.

In practice, local governments make use of a wide array of subsidy measures to provide business incentives to local firms. These include property tax abatements, customized job training subsidies, investment tax credits, research and development tax credits, deal-closing programs, and so on. I do not attempt to directly model all these different policy measures but focus instead on their common effect on business costs. In particular, I simply assume that regional governments offer subsidies to all local firms which pay for a fraction of their overall fixed and variable costs.

This simplification helps me keep the analysis transparent and ensures I model subsidies in a way that is compatible with the aforementioned W.E. Upjohn Institute for Employment Research business incentive database. As I will describe in more detail in the data section, this is the best available database on local business incentives which I use to calibrate the subsidies local governments provide. It aims to measure the "standard deal" available to most medium and medium-large businesses and reports local business incentives as a fraction of local value added.

I interpret these subsidies as deviations from benefit tax rates, i.e. taxes for which firms receive public goods of equal value in return. This allows me to abstract from business taxation and public good provision altogether which further simplifies the analysis. I implement this simplification by interpreting statutory business taxes as benefit taxes which do not affect the location decisions of firms. While this at first looks like a strong assumption, we will see that all results are surprisingly robust to measurement error in the subsidy variable which is where the mistake would show up.

In the end, the only taxes I have in the model are therefore the taxes collected to finance the subsidies. I assume that these taxes are levied on local residents in a lump-sum fashion since they would ultimately have to pay for any shortfall between the revenues from taxes collected from local businesses and the expenditures on public good provision to local businesses. Denoting the proportional subsidy on business costs by s_i, the wage rate by w_i, the interest rate by i, the land rental rate by r_i, and local expenditures on intermediates by E_i^I, the local tax bill is given by:

$$S_i = s_i \left(w_i L_i + i K_i + r_i T_i^C + E_i^I \right) \tag{4}$$

2.1.4 Budget constraint

Local residents earn local labor income $w_i L_i$, local land income $r_i T_i$, and a share of national capital income $\lambda_i^L i K$. $\lambda_i^L \equiv L_i / L$ is simply the share of workers residing in region i so that each worker is assumed to own an equal share of the nation's capital stock. They use this income for their expenditures on final goods E_i^F, residential land $r_i T_i^R$, and taxes S_i, as well as an interregional transfer Ω_i which satisfies $\sum_{i=1}^{R} \Omega_i = 0$. This transfer helps rationalize inter-regional trade imbalances and captures side payments in the cooperative regime. Their budget constraint is therefore given by:

$$w_i L_i + \lambda_i^L iK + r_i T_i = E_i^F + r_i T_i^R + S_i + \Omega_i \tag{5}$$

In particular, it is easy to show that a region's aggregate net exports are given by $NX_i = \left(\lambda_i^K - \lambda_i^L\right) iK + \Omega_i$, where $\lambda_i^K \equiv K_i/K$ is the share of capital employed in region i. As a result, Ω_i can be calibrated to ensure that the predicted NX_i matches the data, as is commonly done in the trade literature. The term $\left(\lambda_i^K - \lambda_i^L\right) iK$ arises because of the earlier assumption that each worker owns an equal share of the nation's capital stock. It implies that there is a difference between the capital income generated by local firms and the one accruing to local residents whenever $\lambda_i^K \neq \lambda_i^L$ which is then mirrored in net exports.

Building on this intuition, Caliendo et al (2014) have recently suggested an alternative way of dealing with aggregate trade imbalances. In particular, they do not assume that each worker owns an equal share of the nation's capital stock but instead make workers' asset holdings dependent on their state of residence. For example, workers in Florida are assumed to own a larger share of the nation's assets which then allows them to finance their state's trade deficit. The authors show that one can calibrate state-specific ownership shares in that manner to largely explain the observed trade deficits.

While I am sympathetic to this idea, I believe it is not well suited for my application because it implies that workers' asset holdings change whenever they switch locations. For example, workers would then benefit from moving to Florida simply because this would give them a larger share in the nation's asset holdings which would clearly distort my policy analysis. In any case, it would also be just a patch for the more fundamental problem that it is hard to rationalize aggregate trade imbalances in static models since they are ultimately driven by intertemporal savings and investment decisions.

2.2 Equilibrium in levels

To set the stage for my analysis of non-cooperative and cooperative subsidies, I begin by characterizing the equilibrium for given subsidies. In this equilibrium, workers maximize utility, firms maximize profits, free entry ensures zero profits, and all goods and factor markets clear. It can be expressed as a system of $4R$ equations in the $4R$ unknowns P_i, λ_i^L, λ_i^K, λ_i^C, where P_i is the price index dual to C_i^F and C_i^I, λ_i^L and λ_i^K are the regional labor and capital employment shares defined earlier, and $\lambda_i^C \equiv T_i^C/T_i$ is the share of land in i used for commercial purposes. In particular:[6]

Definition 1 *For given subsidies and a numeraire $i \equiv 1$, an equilibrium in levels is a set of $\left\{P_i, \lambda_i^L, \lambda_i^K, \lambda_i^C\right\}$ such that*

$$\lambda_i^L = \frac{U_i^\sigma}{\sum_{j=1}^R U_j^\sigma} \tag{6}$$

$$P_j = \left(\sum_{i=1}^R M_i \left(p_{ij}\right)^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}} \tag{7}$$

$$\frac{1}{\varepsilon}\sum_{j=1}^R \left(p_{ij}\right)^{1-\varepsilon}\left(P_j\right)^{\varepsilon-1} E_j = \left(\left(w_i\right)^{\theta^L}\left(r_i\right)^{\theta^T}\right)^\eta \left(P_i\right)^{1-\eta}\rho_i f_i \tag{8}$$

$$r_i T_i = \frac{\mu}{1-\mu}E_i^F + \frac{\eta\theta^T}{1-\eta}E_i^I \tag{9}$$

where

$$w_i = \frac{\lambda_i^K}{\lambda_i^L}\frac{\theta^L}{\theta^K}\frac{K}{L} \tag{10}$$

$$r_i = \frac{\lambda_i^K}{\lambda_i^C}\frac{\theta^T}{\theta^K}\frac{K}{T_i} \tag{11}$$

$$E_i^I = \frac{1-\eta}{\eta\theta^K}\lambda_i^K K \tag{12}$$

$$p_{ij} = \frac{\varepsilon}{\varepsilon-1}\frac{\left(\left(w_i\right)^{\theta^L}\left(r_i\right)^{\theta^T}\right)^\eta \left(P_i\right)^{1-\eta}\rho_i\tau_{ij}}{\varphi_i} \tag{13}$$

$$S_i = s_i\lambda_i^K\frac{iK}{\eta\theta^K} \tag{14}$$

$$\Omega_i = NX_i - \left(\lambda_i^K - \lambda_i^L\right)K \tag{15}$$

$$E_i^F = (1-\mu)\left(w_i L_i + \lambda_i^L K + r_i T_i - \left(S_i + \Omega_i\right)\right) \tag{16}$$

$$E_i = E_i^F + E_i^I \tag{17}$$

[6] In the interest of brevity, I only provide an intuitive discussion of these and all other equations in the main text. I happily provide step by step derivations upon request.

$$U_i = \frac{1}{1-\mu} \frac{A_i}{L_i} \frac{E_i^F}{(r_i)^{\mu} (P_i)^{1-\mu}} \tag{18}$$

$$M_i = \frac{L_i}{\varepsilon f_i \eta \theta^L} \frac{w_i}{\left((w_i)^{\theta^L} (r_i)^{\theta^T}\right)^{\eta} (P_i)^{1-\eta}} \tag{19}$$

This says that equations (6) - (9) can be reduced to a system of $4R$ equations in the $4R$ unknowns P_i, λ_i^L, λ_i^K, and λ_i^C by substituting equations (10) - (19). In particular, equations (10) - (19) can be used to successively solve for their respective left-hand side variables in terms of P_i, λ_i^L, λ_i^K, λ_i^C, and parameters which can then be substituted to eliminate those variables from equations (6) - (9). While this is easy to do, the resulting reduced-form equations become rather cumbersome so that it makes more sense to discuss their underlying intuitions by considering the more transparent building blocks (6) - (19).

Equation (6) follows from the fact that $prob\,(U_{iv} \geq U_{jv} \text{ for all } j \neq i) = \frac{U_i^{\sigma}}{\sum_{j=1}^{R} U_j^{\sigma}}$ from the properties of the Frechet distribution, as is also well known from the discrete choice literature. It simply captures that better regions attract more workers, where "better" refers to the common component of utility. This relationship is the stronger the higher is σ, because a high σ corresponds to a low dispersion in idiosyncratic utilities. This equation also reveals that maximizing U_i is similar to maximizing local employment as already mentioned earlier, at least if R is sufficiently large.

Equations (7) - (9) require less of an explanation, as they are simply a CES price index, a zero-profit condition, and a land market clearing condition, respectively, with p_{ij} denoting the delivered price of a good from region i in region j and $\rho_i \equiv 1 - s_i$. In particular, the CES price index takes the standard form, the zero profit condition requires that operating profits equal subsidized fixed costs, and the land market clearing condition imposes that the total land income in region i is equal to the sum of residential and commercial land expenditure in region i.

The intuitions underlying equations (10) - (13) should also be fairly clear. In particular, equations (10) - (12) follow directly from the nested Cobb-Douglas structure of the production function which implies that firms spend a share $\eta\theta^L$ of their costs on labor, a share $\eta\theta^K$ of their costs on capital, a share $\eta\theta^T$ of their costs on commercial land, and a share $1-\eta$ of their costs on intermediates. Moreover, equation (13) captures that prices are constant markups over subsidized marginal costs, where $\tau_{ij} > 1$ is an iceberg transport cost in the sense that τ_{ij} units need to be shipped from i for 1 unit to arrive in j.

Equation (14) is a compact version of the earlier equation (4) which summarizes subsidy costs. It is obtained by substituting equations (10) - (12) into equation (4) after rewriting equations (10) - (11) in terms of $w_i L_i$ and $r_i T_i^C$ which requires using the earlier definitions $\lambda_i^L = \frac{L_i}{L}$, $\lambda_i^K = \frac{K_i}{K}$, and $\lambda_i^T = \frac{T_i^C}{T_i}$. It says that local subsidy costs are increasing in the local subsidy rate and the share of capital employed locally which effectively serves as a proxy for the size of the subsidized local economy since the local uses of labor, capital, commercial land, and intermediate inputs comove.

Equations (15) - (17) calculate transfers as well as final and overall expenditure on goods. Equation (15) is simply a rearranged version of the earlier relationship $NX_i = \left(\lambda_i^K - \lambda_i^L\right) iK + \Omega_i$, where NX_i is set to match the aggregate net exports of region i. Equation (16) follows from the budget constraint (5) and the fact that consumers spend a share $1 - \mu$ of their income on goods and the remainder on residential land. Equation (17) simply says that total expenditure on goods consists of expenditure on final goods by consumers and intermediate goods by firms.

This leaves me with equations (18) and (19) to explain. Equation (18) is simply amenity adjusted per-capita real income since $\frac{1}{1-\mu} E_i^F$ is total expenditure on residential land and final goods and $(r_i)^\mu (P_i)^{1-\mu}$ is the corresponding aggregate price index. Equation (19) follows from the fact that zero profits imply that firms must be of a constant size $z_i = \varepsilon f_i$, as is typically the case in such environments. This then implies that the number of firms is given by $M_i = \frac{1}{\varepsilon f_i} \left(\frac{1}{\eta} \left(\frac{L_i}{\theta^L}\right)^{\theta^L} \left(\frac{K_i}{\theta^K}\right)^{\theta^K} \left(\frac{T_i^C}{\theta^T}\right)^{\theta^T} \right)^{\eta} \left(\frac{C_i^I}{1-\eta}\right)^{1-\eta}$ which further simplifies to equation (19) upon substituting equations (10) - (12).

2.3 Equilibrium in changes

Before using this system of equations to analyze non-cooperative and cooperative subsidies, it is convenient to first express it in changes following Dekle et al's (2007) "exact hat algebra". This technique is now standard in the quantitative trade literature and has also been applied recently in economic geography settings (see, for example, Redding 2016). Here, the main advantage is that it eliminates the need to explicitly estimate the technology parameters φ_i

and f_i, the preference parameters A_i, and the trade cost parameters τ_{ij}, thereby very much simplifying the quantitative analysis. In particular:

Definition 2 *For given subsidy changes and a numeraire $i \equiv 1$, an equilibrium in changes is a set of $\left\{\hat{P}_i, \hat{\lambda}_i^L, \hat{\lambda}_i^K, \hat{\lambda}_i^C\right\}$ such that*

$$\hat{\lambda}_i^L = \frac{\left(\hat{U}_i\right)^\sigma}{\sum_{j=1}^R \lambda_j^L \left(\hat{U}_j\right)^\sigma} \tag{20}$$

$$\hat{P}_j = \left(\sum_{i=1}^R \alpha_{ij} \hat{M}_i \left(\hat{p}_{ij}\right)^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}} \tag{21}$$

$$\sum_{j=1}^R \beta_{ij} \left(\hat{p}_{ij}\right)^{1-\varepsilon} \left(\hat{P}_j\right)^{\varepsilon-1} \hat{E}_j = \left(\left(\hat{w}_i\right)^{\theta^L} \left(\hat{r}_i\right)^{\theta^T}\right)^\eta \left(\hat{P}_i\right)^{1-\eta} \hat{\rho}_i \tag{22}$$

$$\hat{r}_i = \left(1 - \lambda_i^C\right) \hat{E}_i^F + \lambda_i^C \hat{E}_i^I \tag{23}$$

where

$$\hat{w}_i = \frac{\hat{\lambda}_i^K}{\hat{\lambda}_i^L} \tag{24}$$

$$\hat{r}_i = \frac{\hat{\lambda}_i^K}{\hat{\lambda}_i^C} \tag{25}$$

$$\hat{E}_i^I = \hat{\lambda}_i^K \tag{26}$$

$$\hat{p}_{ij} = \left(\left(\hat{w}_i\right)^{\theta^L} \left(\hat{r}_i\right)^{\theta^T}\right)^\eta \left(\hat{P}_i\right)^{1-\eta} \hat{\rho}_i \tag{27}$$

$$S_i' = s_i' \lambda_i^K \hat{\lambda}_i^K \frac{K}{\eta \theta^K} \tag{28}$$

$$\hat{E}_i^F = (1-\mu) \left(\frac{w_i L_i}{E_i^F} \hat{w}_i \hat{\lambda}_i^L + \lambda_i^L \hat{\lambda}_i^L \frac{K}{E_i^F} + \frac{r_i T_i}{E_i^F} \hat{r}_i - \frac{S_i' + \Omega_i'}{E_i^F}\right) \tag{29}$$

$$\hat{E}_i = \frac{E_i^F}{E_i} \hat{E}_i^F + \frac{E_i^I}{E_i} \hat{E}_i^I \tag{30}$$

$$\hat{U}_i = \frac{1}{\hat{\lambda}_i^L} \frac{\hat{E}_i^F}{\left(\hat{r}_i\right)^\mu \left(\hat{P}_i\right)^{1-\mu}} \tag{31}$$

$$\hat{M}_i = \frac{\hat{w}_i \hat{\lambda}_i^L}{\left((\hat{w}_i)^{\theta^L} (\hat{r}_i)^{\theta^T} \right)^{\eta} \left(\hat{P}_i \right)^{1-\eta}} \tag{32}$$

and

$$\alpha_{ij} = \frac{X_{ij}}{\sum_{m=1}^{R} X_{mj}} \tag{33}$$

$$\beta_{ij} = \frac{X_{ij}}{\sum_{n=1}^{R} X_{in}} \tag{34}$$

$$w_i L_i = \frac{\eta \theta^L}{\rho_i} \sum_{n=1}^{R} X_{in} \tag{35}$$

$$K_i = \frac{\eta \theta^K}{\rho_i} \sum_{n=1}^{R} X_{in} \tag{36}$$

$$r_i T_i^C = \frac{\eta \theta^T}{\rho_i} \sum_{n} X_{in} \tag{37}$$

$$E_i^I = \frac{1-\eta}{\rho_i} \sum_{n=1}^{R} X_{in} \tag{38}$$

$$E_i^F = \sum_{m=1}^{R} X_{mi} - E_i^I \tag{39}$$

$$E_i = E_i^F + E_i^I \tag{40}$$

$$r_i T_i = \frac{\mu}{1-\mu} E_i^F + r_i T_i^C \tag{41}$$

$$\lambda_i^K = \frac{K_i}{\sum_{i=1}^{R} K_i} \tag{42}$$

$$\lambda_i^C = \frac{r_i T_i^C}{r_i T_i} \tag{43}$$

Conditions (20) - (32) are calculated by expressing conditions (6) - (19) in changes, where a "hat" denotes the proportional change of a variable from some original value x to some new value x', $\hat{x} = \frac{x'}{x}$ induced by a change in subsidies (from s_i to s_i') or transfers (from Ω_i to Ω_i'). Using conditions (33) - (43), their coefficients can be expressed in terms of easily observable quantities such as the value of trade flowing from region i to region j, $X_{ij} = M_i (p_{ij})^{1-\varepsilon} (P_j)^{\varepsilon-1} E_j$. In the end, all one needs to solve the model in changes is data on X_{ij}, λ_i^L, and s_i, as well as estimates of the parameters σ, μ, ε, θ^L, θ^K, θ^T, and η.

Besides substantially simplifying the quantification, this exact hat algebra approach also ensures that all counterfactuals are computed from a benchmark which perfectly matches observed regional employment, regional production, regional subsidies, and interregional trade. Essentially, it imposes a restriction on the set of unknown parameters $\{\varphi_i, f_i, A_i, \tau_{ij}\}$ such that the predicted λ_i^L and X_{ij} exactly match the observed λ_i^L and X_{ij} given the observed s_i and the model parameters $\{\sigma, \mu, \varepsilon, \theta^L, \theta^K, \theta^T, \eta\}$. I will elaborate further on this in a later section in which I discuss the model fit.

2.4 Isomorphism

Building on Allen and Arkolakis (2014), I show in Appendix 1 that the model can also be interpreted as an Armington model with external increasing returns to scale. In particular, suppose instead that each region makes one differentiated variety under conditions of perfect competition subject to the aggregate production function $Q_i = \varphi_i (Z_i)^{1+\phi}$, where outputs, Q_i, and inputs, Z_i, are now represented in capital letters to emphasize that they refer to aggregate quantities. $\phi_i > 0$ is an external increasing returns parameter which captures that local productivity is increasing in local employment.

Keeping the rest of the model unmodified, I show in the appendix that such an Armington model is isomorphic to the above New Economic Geography model under the assumption that $\phi = 1/(\varepsilon - 1)$. Intuitively, the local price index is decreasing in local employment in both models, with the mechanism operating through changes in variety in the New Economic Geography model and through changes in productivity in the Armington model. I exploit this feature to assess how robust my results are to my particular model specification by allowing for $\phi \neq 1/(\varepsilon - 1)$ in sensitivity checks.

3 Calibration

3.1 Data

I apply this model to analyze subsidy competition among U.S. states, focusing on manufacturing in the lower 48 states in the year 2007. Recall from the above discussion that I need data on interregional trade flows X_{ij}, employment shares λ_i^L, and subsidies s_i, as well as estimates of the parameters $\{\sigma, \mu, \varepsilon, \theta^L, \theta^K, \theta^T, \eta\}$. I obtain this information from the 2007 Commodity Flow Survey, the 2007 Annual Survey of Manufacturing, the business incentive databases of Bartik (2017) and Story et al (2012), the 2007 BEA Input-Output Table and BLS Capital Income Table, as well as work by Redding (2016) and Suarez Serrato and Zidar (2016).

I construct the matrix of interstate trade flows from the Commodity Flow Survey scaled to match state-level manufacturing production from the Annual Survey of Manufacturing. Using the publicly available Commodity Flow Survey data, I begin by constructing a matrix of interstate freight shipments. I use the reported values which aggregate over all modes of transport and all included industries in order to avoid having to deal with the many missing values there are at finer levels of detail. In the end, there are still about 8 percent missing values, all pertaining to interstate rather than intrastate flows.

I interpolate these missing interstate flows using the standard gravity equation my model implies: $X_{ij} = M_i \left(p_{ii}\tau_{ij}\right)^{1-\varepsilon} \left(P_j\right)^{\varepsilon-1} E_j$. In particular, I estimate this equation by regressing log trade flows on origin fixed effects, destination fixed effects, and standard proxies for trade costs, namely log distance between state capitals and a dummy for whether i and j share a state border. Reassuringly, the estimation delivers a positive common border coefficient and a plausible distance elasticity of trade flows of -1.01. The correlation between predicted values and observed values is 96 percent.

I then scale these freight shipments to ensure they add up to the total manufacturing shipments reported in the Annual Survey of Manufacturing for each state. On average, the total freight shipments implied by the Commodity Flow Survey are almost 2.5 times larger than the total manufacturing shipments reported in the Annual Survey of Manufacturing.[7] However, notice that trade shares and not trade flows enter into equations (20) - (43) used to calculate the effects of subsidy changes so that these scalings only matter if they affect different states differentially.

I obtain the vector of labor shares λ_i^L from the Annual Survey of Manufacturing. In particular, I simply calculate the total number of U.S. manufacturing workers and determine

[7] In part, this simply reflects the fact that the aggregate freight shipments I use from the Commodity Flow Survey include all goods captured by the Standard Classification of Transported Goods which includes not just manufacturing goods. However, the Commodity Flow Survey also double-counts trade flows if they are shipped indirectly, say first from i to m and then from m to j.

the share of those employed in a particular state. These shares range from 0.03 percent for Wyoming to 10.98 percent for California and their distribution is as one would expect. In particular, manufacturing is mainly concentrated in California, Texas, and the traditional manufacturing belt states stretching all the way from New York to Illinois. Also, there is generally little manufacturing activity in the Interior West of the country.

I obtain most of my subsidy measures from a new Panel Database on Business Incentives from the W.E. Upjohn Institute for Employment Research. This database is the best available database on local business incentives and is documented in detail in Bartik (2017). It aims to measure the "standard deal" available to most medium and medium-large businesses and reports local business incentives as a fraction of local value added. It mainly includes property tax abatements, customized job training subsidies, investment tax credits, and research and development tax credits.

As my subsidy measure I use the firm-age weighted average of the present value of business incentives available to manufacturing firms in 2007. This average is calculated by first simulating the present value of business incentives available to new firms over a 20-year time period and then adjusting the starting years to match the firm age distribution in 2007. This measure is readily available from the database and strikes me as the most reasonable one for my purposes since I cannot capture the time profile of business incentives in my static environment.

I supplement the information available from the Panel Database on Business Incentives with information available from the New York Times' Business Incentive Database compiled by Story et al (2012). This is necessary because the Panel Database on Business Incentives currently covers only 32 states plus the District of Columbia in an effort to economize on resources. However, the missing 14 states only account for less than 10 percent of all U.S. private sector GDP so that the gap in the Panel Database on Business Incentives is smaller than it first seems.

In contrast to the Panel Database on Business Incentives, the New York Times' Business Incentive Database does not attempt to back out the "standard deal" available to most businesses but simply reports an estimate of the total annual value of all business incentives including sales tax abatements, property tax abatements, corporate tax abatements, cash grants, loans, and free services. I correct for this discrepancy by scaling the entire New York Times data such that it lines up with the Panel Database on Business Incentives for the 32 states included in both datasets.

Unfortunately, the value of subsidies going to manufacturing firms is not straightforward to determine in the New York Times' Business Incentive Database since many incentive programs are not classified by industry. To obtain at least a rough estimate, I take the value of subsidies going explicitly to manufacturing (around 32 percent), disregard all subsidies going explicitly to agriculture, oil, gas and mining, and film and allocate the residual (about 53 percent) to manufacturing based on manufacturing shares in state GDP obtained from the Bureau of Economic Analysis.

In order to bring these subsidy measures in line with their representation in the theory, I express them as a fraction of total revenues which is the same as total costs since free entry is assumed to drive profits down to zero. The resulting subsidy rates do not exhibit any clear geographic pattern and average 0.5 percent nationwide. New Mexico (3.8 percent), Vermont (3.2 percent), and Oklahoma (2.5 percent) are the three most generous states while Colorado (0.0 percent), Arkansas (0.0 percent), and Delaware (0.0 percent) are the three least generous states.

I estimate the shares of labor, capital, and land in value added from the 2007 input-output tables of the Bureau of Economic Analysis. In particular, I calculate the share of labor in value added as the share of employee compensation in value added net of taxes. I then divide the residual into the capital share and the land share by using the shares of equipment, intellectual property, and inventories in all assets and the share of structures and land in all assets from the 2007 capital income tables of the Bureau of Labor Statistics. Aggregating over all manufacturing industries, I find $\theta^L = 0.57$, $\theta^K = 0.33$, and $\theta^T = 0.10$.

I use the same input-output tables to calculate the share of value added in gross production. In doing so, I have to recognize that my model does not directly map into published input-output tables for two reasons. First, I do not have any investment in my model while the published input-output tables distinguish between purchases which are depreciated immediately and purchases which are capitalized on the balance sheet. Second, I only have manufacturing industries in my model while the published input-output tables encompass the entire economy.

I deal with the first issue by scaling all rows in the main body of the use table by one plus the ratio of private fixed investment to total intermediates. By doing so, I effectively treat all purchases firms make as intermediate consumption which matters mostly for durable goods industries such as machinery. Otherwise, I would essentially assume that firms do not value cheap access to machinery only because they capitalize them on their balance sheets. I deal with the second issue by simply cropping the input-output table to include only manufacturing industries. Using this procedure, I find $\eta = 0.58$.

I take the remaining parameters μ, σ, and ε from the literature. I particular, I set $\mu = 0.25$ following Redding (2016) who bases his choice on housing expenditure shares documented by Davis and Ortalo-Magne (2011). Moreover, I set $\sigma = 1.2$ as in Suarez Serrato and Zidar (2016) who estimate it by exploiting the fact that it also represents a local labor supply elasticity. Finally, I pick a value of $\varepsilon = 5$ which represents a typical estimate from the trade literature. Needless to say, the estimates of σ and ε have to be handled with particular caution so I also provide extensive sensitivity checks.

3.2 Adjustments

As laid out so far, the framework has two debatable implications which I will now discuss. First, subsidies can have an efficiency enhancing effect in addition to their main beggar-thy-neighbor effect since goods prices are too high relative to land and factor prices as a result of a markup distortion. Second, subsidies can have a second beggar-thy-neighbor effect in addition to their main agglomeration effect since they also bring about an interregional wealth redistribution by affecting the real value of the nominal transfers which were introduced to rationalize aggregate trade deficits.

It is not clear how to best deal with the issue that subsidies can have an efficiency enhancing effect. Essentially, one can either eliminate the markup distortion or embrace it as a central feature of the economic environment. The former approach can be justified by arguing that the markup distortion is just one of many distortions affecting real-world economies and therefore should not be overemphasized. The latter approach can be defended by pointing out that the markup distortion is not just any distortion but one that is intimately related to the agglomeration externality.

The intimate relationship between the allocative inefficiency and the agglomeration externality is particularly clear in the isomorphic external increasing returns to scale representation introduced above. In this representation, the external increasing returns not only allow regions to gain at the expense of one another but also imply that goods are underprovided due to a wedge between private and social marginal costs. This implies that the same local spillovers which make subsidies beggar-thy neighbor policies also bring about the allocative inefficiency which subsidies can correct.

In light of this, I report results following both approaches so that readers can make their own choice. In particular, I extend the model by allowing for a federal cost subsidy financed by lump-sum taxes on all national residents. This federal subsidy is set to exactly neutralize the markup distortion so that state subsidies then have no additional efficiency enhancing effect (the details can be found in Appendix 2). When I discuss my findings, I always start by considering the case with such a federal subsidy and then ask how the results change if it is removed.

The prediction that subsidies also bring about an interregional wealth redistribution strikes me as collateral damage from a modeling patch. The issue is simply that the nominal transfer Ω_j is evaluated in real terms in the indirect utility function so that $\frac{\Omega_j}{P_j}$ is what governments care about. One implication of this is that governments then have an incentive to manipulate relative prices such that the real value of the transfer they make (receive) is minimized (maximized). Unfortunately, this incentive is strong enough to severely contaminate the quantitative results given the large trade imbalances in the dataset.

In order to avoid this problem, I follow my approach in Ossa (2014) and first use the model to purge the trade data from the interregional transfers and then work with the purged data subsequently. Notice that this could be done by setting $\Omega_i' = 0$ and $s_i' = s_i$ in equations (20) - (43) and then calculating the implied trade flows using $\hat{X}_{ij} = \hat{M}_i \left(\hat{p}_{ii}\right)^{1-\varepsilon} \left(\hat{P}_j\right)^{\varepsilon-1} \hat{E}_j$. However, I use a slightly modified version of the model in an attempt to minimize the difference between the purged data and the original data. In particular, I treat λ_i^L as exogenous by setting $\hat{\lambda}_i^L = 1$ and dropping equation (20).[8]

[8]Another advantage of purging the data from interregional transfers is that I do not have to take a stance

This procedure does not affect the pattern of interregional trade flows with the correlation between original and purged data being 99.1 percent. It also does not affect the cross-regional distribution of capital with the correlation between original and purged capital shares λ_i^K being 99.9 percent (recall that the labor shares λ_i^L are held fixed). Just as in Dekle et al (2007), the main effect is that the prices of fixed factors rise (fall) in regions running trade surpluses (deficits). The adjustments in $\hat{w}_i^{\theta^L} \hat{r}_i^{\theta^T}$ range from -18.1 percent in Montana to 6.0 percent in Wisconsin and are between -5.5 percent and 6.0 percent for 44 out of 48 states.

3.3 Multiplicity

As is usually the case in New Economic Geography models, there are multiple equilibria if the agglomeration forces are sufficiently strong relative to the dispersion forces. Concretely, this means that equations (20) - (43) can have solutions other than $\hat{P}_i = \hat{\lambda}_i^L = \hat{\lambda}_i^K = \hat{\lambda}_i^C = 1$ for factual subsidies, which is always an equilibrium by construction because it corresponds to the factual situation. Multiple equilibria are more likely the higher is σ since location preferences are then less dispersed. Multiple equilibria are also more likely the lower is ε since consumers and firms then care more about being close to firms.

Figure 1 illustrates the structure of equilibria in the calibrated model for various values of σ and ε. This figure is constructed by checking if equations (20) - (43) converge to different solutions for a large sequence of random starting guesses over a fine grid of values for σ and ε. As can be seen, my benchmark values $\sigma = 1.2$ and $\varepsilon = 5$ are safely within the region in which there is a unique equilibrium. The same is true for all values within the ranges $\sigma \in [0.8, 1.6]$ and $\varepsilon \in [4, 6]$ which I work with in sensitivity checks (labelled "range of considered parameters" in the figure).

3.4 Model fit

Model fit is typically not discussed in papers using Dekle et al's (2007) "exact hat algebra" method since the model perfectly fits the data used in the calibration by construction. This is no different in my application, where the method essentially imposes a restriction on the set

on the units in which they are held fixed. This would raise serious interpretational issues which are usually ignored in the quantitative trade literature.

of unknown parameters $\{\varphi_i, f_i, A_i, \tau_{ij}\}$ such that the predicted λ_i^L and X_{ij} exactly match the observed λ_i^L and X_{ij} given the observed s_i and the model parameters $\{\sigma, \mu, \varepsilon, \theta^L, \theta^K, \theta^T, \eta\}$. As is usually the case, the unknown parameters are not uniquely identified since there are more parameters than empirical moments.

However, some progress can be made by imposing the restrictions $\tau_{ij} = \tau_{ji}$ and $\tau_{ii} = 1$ for all i and j. In particular, it is then possible to "invert" the model and back out relative trade costs, amenities, productivities, and many other variables which seems useful to get a sense of the parameter variation needed to explain the observed economic geography. As I discuss in more detail in Appendix 3, the variation in trade flows is mainly explained by variation in trade costs which are highly correlated with distance. Moreover, the variation in manufacturing employment is mainly explained by variation in amenities with Wyoming and California having the worst and best amenities, respectively.

4 Analysis

4.1 Welfare effects of subsidies

Figure 2 summarizes what happens if Illinois unilaterally deviates from its factual subsidy indicated by the vertical line. The top panel depicts Illinois' local welfare change as well as the average of the local welfare changes of all other states. The center panel shows the change in the number of firms in Illinois as well as the average of the changes in the number of firms in all other states. The bottom panel summarizes the effects on the shares of labor and capital employed in Illinois. As can be seen, higher subsidies allow Illinois to gain at the expense of other states and attract firms, labor, and capital to Illinois.[9]

These local welfare changes are driven by a combination of home market effects, terms-of-trade effects, and congestion effects. Essentially, Illinois gains by attracting economic activity from other states because this reduces Illinois' price index and improves Illinois' terms-of-trade. These gains are attenuated by competition for Illinois' fixed factor land which drives up Illinois' land rental rates and subjects Illinois' firms to diminishing returns. This can be

[9] I started this project when I was still at the University of Chicago which is why I always use Illinois as an example. There is nothing special about Illinois and I could have used any other state.

best explained with reference to the following decomposition of the welfare effects of small subsidy changes which I derive in the appendix:

$$\begin{aligned}\frac{dU_j}{U_j} &= \underbrace{\frac{1}{\eta}\sum_i \alpha_{ij}\frac{1}{\varepsilon-1}\frac{dM_i}{M_i}}_{\text{home market effect}} + \underbrace{\frac{1}{\eta}\sum_i \alpha_{ij}\left(\frac{dp_{jj}}{p_{jj}}-\frac{dp_{ii}}{p_{ii}}\right)}_{\text{terms-of-trade effect}} \\ &\quad - \underbrace{\mu\left(\frac{dr_j}{r_j}-\frac{dP_j}{P_j}\right)}_{\text{residential congestion}} - \underbrace{\theta^T\left(\frac{d\lambda_j^L}{\lambda_j^L}-\frac{d\lambda_j^C}{\lambda_j^C}\right)}_{\text{commercial congestion}}\end{aligned} \tag{44}$$

The first term captures a home market effect which is also sometimes referred to as firm relocation or firm delocation effect. In particular, Illinois' subsidy induces some firms to relocate to Illinois from other states. This has two conflicting effects on Illinois' price index since Illinois' consumers now have access to more domestic varieties but fewer foreign varieties. However, Illinois' consumers gain more from the increase in the number of domestic varieties than they lose from the decrease in the number of foreign varieties since they spend more on domestic varieties because of trade costs.

The second term captures a terms-of-trade effect. In particular, the relocation of economic activity to Illinois increases labor and land demand in Illinois relative to other states so that Illinois' wage and land rental rates increase relative to other states. Given that wage and land rental rate changes directly translate into price changes in this constant markup environment, this then increases the prices of goods Illinois exports to other states relative to the prices of goods Illinois imports from other states which amounts to an improvement in Illinois' terms-of-trade.

While these relative wage and relative rent effects are the dominant effects on Illinois' terms-of-trade, two additional effects need to be taken into account. In particular, there is an adverse direct subsidy effect which arises because Illinois' subsidies directly reduce the price of goods made in Illinois. Also, there is an adverse intermediate cost effect which arises because production relocations to Illinois reduce the price index of intermediate goods in Illinois. Defining $\frac{dToT_j}{ToT_j} = \frac{1}{\eta}\sum_i \alpha_{ij}\left(\frac{dp_j}{p_j}-\frac{dp_i}{p_i}\right)$, this can be seen immediately from the pricing equation (13) which implies:

$$\frac{dToT_j}{ToT_j} = \underbrace{\theta^L \sum_i \alpha_{ij} \left(\frac{dw_j}{w_j} - \frac{dw_i}{w_i} \right)}_{\text{relative wage effect}} + \underbrace{\theta^T \sum_i \alpha_{ij} \left(\frac{dr_j}{r_j} - \frac{dr_i}{r_i} \right)}_{\text{relative rent effect}} \tag{45}$$
$$+ \underbrace{\frac{1}{\eta} \sum_i \alpha_{ij} \left(\frac{d\rho_j}{\rho_j} - \frac{d\rho_i}{\rho_i} \right)}_{\text{direct subsidy effect}} + \underbrace{\frac{1-\eta}{\eta} \sum_i \alpha_{ij} \left(\frac{dP_j}{P_j} - \frac{dP_i}{P_i} \right)}_{\text{intermediate cost effect}}$$

The third and fourth terms in equation (44) capture residential and commercial congestion effects, respectively. Residential congestion arises if local workers consume less residential land relative to final goods which happens if local land rental rates rise relative to the local goods price index. Commercial congestion arises if local firms run into diminishing returns by using less commercial land per worker which happens if local land rental rates rise faster than local wages. The latter point follows from the fact that $\frac{d\lambda_i^L}{\lambda_i^L} - \frac{d\lambda_i^C}{\lambda_i^C} = \frac{dr_i}{r_i} - \frac{dw_i}{w_i}$, which should be straightforward to verify.

As an illustration, I have used formula (44) to decompose the effects of a 5 percent subsidy imposed by Illinois. Illinois' welfare goes up by 1.2 percent of which 1.6 percent are due to home market effects, 1.0 percent are due to terms-of-trade effects, and -1.4 percent are due to congestion effects. The terms-of-trade effect consists of relative wage, relative rent, direct subsidy, and intermediate cost effects of 5.4 percent, 0.5 percent, -4.5 percent, and -0.3 percent, respectively. The congestion effect consists of residential and commercial congestion effects of -2.3 percent and 0.7 percent, respectively.[10]

4.2 Optimal subsidies

I now compute the optimal subsidies of all 48 states, assuming each time that all other states do not deviate from their factual subsidies. The goal is to quantify how much states could gain from unilateral policy interventions and set the stage for the subsequent analysis of subsidy wars. As I describe in detail in Appendix 4, I compute optimal subsidies by maximizing G_j^{loc} as defined in equation (3) using the Su and Judd (2012) method of mathematical programming

[10] I have scaled all effects from decomposition (44) so that they sum to the welfare effects computed using the system of equilibrium conditions. This is necessary because equation (44) is just a linear approximation for discrete subsidy changes which ignores second-order distortions to expenditure shares.

with equilibrium constraints. This ensures fast convergence despite the high dimensionality of the analysis.

Figure 3 summarizes the optimal subsidies of all 48 states. As can be seen, they range from 5.8 percent for Tennessee to 12.2 percent in Louisiana and are strongly related to states' own trade shares. The own trade share is an inverse measure of a state's trade openness calculated as the share of purchases it makes from itself. The variation in the own trade shares is mainly driven by variation in trade costs even though state size of course also plays a role. For example, California has by far the highest own trade share and also by far the largest manufacturing employment share.

The tight optimal subsidy-own trade share relationship can be explained with reference to the home market effect which is the dominant effect throughout the analysis. In particular, recall that consumers gain more from the larger number of domestic firms than they lose from the smaller number of foreign firms because they spend more on domestic varieties than on foreign varieties. The own trade share essentially quantifies how much more they spend on domestic varieties than on foreign varieties and therefore determines how much they gain from attracting firms.

Figure 4 turns to the local welfare effects associated with the optimal subsidies from Figure 3. As can be seen, they range from 0.2 percent for New Mexico to 4.6 percent for California and are strongly increasing in states' optimal subsidies. Recall from above that New Mexico, Vermont, and Oklahoma have the highest factual subsidy rates in my sample which explains why they appear as an outlier in Figure 4. In particular, all welfare effects are measured relative to factual subsidies and not zero subsidies and these three states are already closer to their optimal subsidies in the baseline case.

Figure 5 confirms the earlier claim that optimal subsidies are very close to employment-maximizing subsidies. One way to interpret this is that the results are robust to governments maximizing local employment instead of local welfare. Another way to interpret this is that local employment maximization is a good rule of thumb for local welfare maximization. Either way, it is a comforting finding since local jobs feature most prominently in real-world policy debates. It arises simply because workers move to the states which are most attractive as captured by the relationship $\lambda_i^L = \frac{U_i^\sigma}{\sum_{j=1}^R U_j^\sigma}$.

Figure 6 illustrates the geographic propagation of the local welfare effects of optimal subsides using again the example of Illinois. It shows that most of Illinois' neighbors actually gain from Illinois' optimal subsidies with the losses arising in more distant states. The reason is simply that Illinois' neighbors trade a lot with Illinois and can therefore reap some of the benefits of Illinois' increased product variety. While this may be obvious in the context of this model, it does not always seem to be appreciated by real world policymakers who sometimes worry particularly about subsidies imposed by neighboring states.

Table 1 expands on the results from Figures 3 and 4 listing the optimal subsidies together with their welfare effects. It reports the optimal subsidies as well as the local welfare gains of the subsidy imposing state (under "own"), the average local welfare losses in all other states (under "other"), and the national welfare loss (under "national"). The optimal subsidies and local welfare effects are also reported in dollar terms, where the dollar values are calculated by multiplying subsidy rates with subsidy bases and percentage local welfare changes with local per-capita final expenditures.[11]

Optimal subsidies average 9.6 percent or \$14.9 billion, would raise local welfare by an average 2.2 percent or \$1.2 billion in the subsidy imposing state, and would lower local welfare by an average -0.2 percent or -\$2.9 billion in all other states. Notice that the dollar gains of the subsidy imposing state are always smaller than the dollar losses of all other states combined which suggests that subsidies are an inefficient beggar-thy-neighbor policy. This is then also corroborated by the national welfare effects in Table 1 which are all negative and average -0.07 percent.

Table 2 shows the results of four sensitivity checks. Panel A reports the sensitivity of the results to the value of σ within roughly the 95 percent confidence interval of the estimate reported by Suarez Serrato and Zidar (2016). Recall that σ is an inverse measure of the dispersion of workers' location preferences so that a higher σ means that workers are more willing to move. For each value of σ, Panel A reports the average optimal subsidy and average

[11] Recall from above that I refer to changes in U_j as local welfare changes and changes in $\left(\sum_{i=1}^R U_i^\sigma\right)^{\frac{1}{\sigma}}$ as national welfare changes and that subsidy induced changes in local welfare correspond to changes in local per-capita real income. By multiplying the percentage local welfare changes with local per-capita final expenditures I obtain the dollar changes which correspond to the percentage changes for fixed prices.

own, other, and national welfare effect analogously to the last line in Table 1. As can be seen, all result are remarkably robust to variation in σ.

Panel B considers the sensitivity of varying ε following the same format as Panel A. ε is the elasticity of substitution among product varieties for which the trade literature has identified $[4, 6]$ as a reasonable range. As can be seen, the optimal subsidies and their welfare effects are strongly decreasing in ε which makes sense since ε is also an inverse measure of the agglomeration externality. This perhaps most obvious in the isomorphic external increasing returns model introduced earlier in which $\phi = 1/(\varepsilon - 1)$ parametrizes the strength of the external increasing returns.

Panel C turns to the sensitivity of varying ϕ in the external increasing returns model now keeping ε unchanged. In this case, the New Economic Geography model and the external increasing returns model are no longer isomorphic so that we can assess what role the particular model specification plays. To make Panels A and B comparable, I calculate the range of ϕ in Panel C by applying the formula $\phi = 1/(\varepsilon - 1)$ to the range of ε in Panel B. As can be seen, the optimal subsidies and their welfare effects are again strongly increasing in ϕ, now even more so than implicit in Panel B.

Panel D suggests that measurement error in my subsidy dataset would only have minimal effects on the results. This is important since I interpret subsidies as deviations from benefit tax rates in the theory which does not map exactly into the measured subsidy rates. In particular, Panel D shows the maximum and minimum optimal subsidies I obtain in 1,000 calculations in which I replace the measured subsidy rates with a bootstrap sample. These maximum and minimum values are very similar in all cases which implies that the optimal subsidies do not depend much on the measured subsidies.

Figures 7 and 8 explore the effects of removing the federal subsidy which was imposed to correct for the markup distortion so far. As can be seen, the optimal subsidies become a bit larger and their "own" welfare effects become a bit smaller while their overall pattern is preserved. On average, the optimal subsidy is 10.1 percent with "own" and "other" welfare effects of 1.5 percent and -0.03 percent, respectively. As I discussed earlier, state subsidies then also have an efficiency enhancing character in addition to their beggar-thy-neighbor character because they counteract the markups charged by firms.

4.3 Nash subsidies

I now turn to the best-response equilibrium in which all states retaliate optimally. This is meant to capture the extreme case of fully non-cooperative policy making which I will also refer to as a subsidy war. As I explain in detail in Appendix 4, it can be found by iterating over the algorithm used to compute optimal subsidies until a fixed point is reached. I have experimented extensively with this procedure and it appears that the fixed point is unique. To avoid confusion, I call the resulting best-response subsidies Nash subsidies and continue using the term optimal subsidies as before.

Figure 9 plots the Nash subsidies against the optimal subsidies from Figure 3. As can be seen, the Nash subsidies tend to be slightly lower than the optimal subsidies but the overall correlation is very high. Intuitively, optimal subsidies are higher than Nash subsidies because states' own trade shares respond more to optimal subsidies than to Nash subsidies. In particular, states attract more firms if other states do not retaliate which then induces them to spend more on domestic goods. This, in turn, magnifies states' incentives to impose further subsidies following the logic discussed above.

Figures 10 and 11 illustrate the local welfare effects associated with these Nash subsidies. As can be seen, they range from -3.1 percent for Delaware to 2.3 percent in Montana so that not all states lose from a subsidy war. The welfare gains are decreasing in states' capital-labor ratios since subsidies are financed by workers but cover labor and capital costs. As a result, per-capita taxes are higher in capital abundant states which results in lower per-capita real incomes there. At the same time, subsidies are no longer effective at inducing firm relocations because they are offered everywhere.

Table 3 elaborates on these figures analogously to Table 1. Of course, there is now only one set of Nash subsidies instead of 48 sets of optimal subsidies so that there is no distinction between "own" and "other" welfare effects. On average, Nash subsidies are 9.1 percent or \$9.9 billion and bring about local welfare losses of -1.1 percent or -\$0.6 billion. These local welfare losses add up to -\$30.9 billion across the entire country and the national welfare loss is -1.3 percent. All in all, a full-out escalation of subsidy competition would therefore have large negative welfare effects.

Table 4 considers the sensitivity of these findings analogously to Table 2. In particular, its various panels again report the effects of changing the parameters σ, ε, and ϕ in the New Economic Geography or external increasing returns version of the model as well as the minimum and maximum Nash subsidies obtained when by replacing the subsidy data with a bootstrap sample 1,000 times. Just as in the case of optimal subsidies, the Nash subsidy results are very robust to changes in σ or measurement error in the subsidy data but strongly respond to changes in ε and ϕ.

Figures 12 and 13 explore the effects of removing the federal subsidy analogously to Figures 7 and 8. While the Nash subsidies are again rather similar with and without the federal subsidy, it turns out that a subsidy war benefits all states without the federal subsidy. As should be clear by now, the reason is that the state subsidies counteract the markup distortion which consumers of intermediate and final goods otherwise face. Essentially, states then unintentionally improve the efficiency of the national economy as their attempts to attract firms from each other more or less cancel out.

My assessment of the welfare effects of subsidy wars therefore critically depends on whether or not I start from a first-best or a laissez-faire benchmark. As I explained earlier, there are good reasons for making either comparison so that I hesitate to take a strong stance. What is clear, however, is that subsidy wars at best move the economy in the right direction and are not a substitute for first-best policies. This is also why I emphasize the case with federal subsidies in most of the paper because I do not want to mislead the reader into endorsing distortionary policies.

4.4 Cooperative subsidies

I now consider cooperative subsidy policy leaving behind the best-response logic from the subsidy war. The goal is to characterize the best-case scenario and assess how much there is to gain relative to the status quo. I assume that the federal government sets state subsidies as well as interstate transfers Ω_j with the objective of maximizing national welfare. As I explain in detail in Appendix 4, I again use the Su and Judd (2012) method of mathematical programming with equilibrium constraints which ensures fast convergence despite the high dimensionality of the analysis.

As one would expect, the cooperative state subsidies are zero or such that prices get reduced by the extent of the markup, depending on whether or not the federal government already corrects for the markup distortion with a federal subsidy. In this case, there is no meaningful distinction between either scenario because a common federal subsidy or uniform state subsidies achieve exactly the same policy goal. As is illustrated in Figure 14, the interstate transfers are used to redistribute per capita income with the result of reducing but not eliminating interstate inequality.

This redistribution improves national welfare by allowing more workers to live in states that better match their idiosyncratic preferences. In particular, some workers in richer states are attracted purely by better consumption possibilities in the sense that their u_{jv}'s are actually higher for poorer states. Transfers from richer to poorer states allow some of these workers to relocate to states for which they have higher u_{jv}'s thereby improving the average match quality. At the same time, there is still inequality in the cooperative equilibrium since transfers come at the cost of reducing production efficiency.[12]

Starting at factual subsidies, cooperation would increase national welfare by 0.5 percent. Almost the entire effect is due to the use of transfers, setting subsidies to zero alone only brings about a welfare gain of 0.002 percent. From a welfare perspective, factual subsidies are therefore much closer to the best-case scenario than the worst-case scenario (recall that the national welfare loss of moving to Nash subsidies is 1.3 percent). Table 5 provides more detail on these numbers which Table 6 complements with sensitivity checks analogous to the earlier Tables 1-4.[13]

Figure 15 compares non-cooperative and cooperative subsidies to factual subsidies. The light grey lines represent non-cooperative subsidies at various degrees of escalation. In particular, the top grey line shows the fully non-cooperative subsidies while the bottom grey line (the x-axis) shows the fully cooperative subsidies and the intermediate grey lines show

[12]If I did not allow the federal government to set interstate transfers, it would attempt to achieve a similar redistribution by manipulating the terms-of-trade using state subsidies. In particular, it would set higher subsidies in poorer states than in richer states thereby improving the terms-of-trade of poorer states relative to richer states.

[13]I do not report the sensitivity of cooperative subsidies with respect to initial subsidies because cooperative subsidies are always zero anyway. Careful readers might notice that there are minor deviations from zero in two of the reported sensitivity checks (for $\sigma = 1$ and $\phi = 0.20$) which I believe are due to computational imprecisions.

proportionately scaled versions with the scalings decreasing in 10 percentage point steps. The factual subsidies are superimposed onto this using state labels and states are always sorted in increasing order of their non-cooperative subsidies.

With the exception of a few outliers, factual subsidies are much closer to cooperative than non-cooperative subsidies. This is not surprising since one would not expect U.S. states to be in a fully escalated subsidy war. Besides perhaps engaging in tacit cooperation, U.S. states also act in the shadow of the federal government which might try to restrict subsidy competition if it became too extreme. For example, the federal government could adopt the argument of some legal scholars that state incentive programs violate the constitution's Commerce Clause because they discriminate against out-of-state businesses.[14]

Figure 16 compares (log) subsidy costs in the factual equilibrium to (log) subsidy costs in the Nash equilibrium. It shows that states with higher factual subsidies also tend to have higher Nash subsidies, which is less apparent when subsidies are expressed as percentage rates in Figure 15. While this is an encouraging observation, it clearly has to be taken with a large grain of salt. Most importantly, the factual subsidy costs I measure are likely to be incomplete and imprecise proxies for the business incentives state governments actually provide, as I discussed in the data section above.

Brushing measurement concerns aside for a moment, one can actually make optimal subsidies line up exactly with factual subsidies by allowing state governments to be partially cooperative. In particular, suppose that state governments maximize a Cobb-Douglas combination of local welfare and national welfare with local welfare weights ν_j: $(U_j)^{\nu_j}\left(\sum_{i=1}^{R} U_i^{\sigma}\right)^{\frac{1-\nu_j}{\sigma}}$. The local welfare weights listed in Table 7 then equalize optimal subsidies and factual subsidies as is verified in Figure 17. Notice that these weights are all below 1 percent suggesting again that the factual regime is close to cooperative.[15]

[14] More precisely, the argument refers to the "dormant" Commerce Clause which U.S. courts have inferred from the Commerce Clause of the U.S. constitution. It holds that states are prohibited from passing legislation which interferes with interstate commerce even if Congress does not intervene. The legal debate therefore focuses on the question of whether state incentive programs interfere with interstate commerce. See Rogers (2000) for an interesting overview.

[15] Given the high correlation between optimal subsidies and Nash subsidies, these weights also bring Nash subsidies close to factual subsidies.

5 Conclusion

In this paper, I provided a first comprehensive quantitative analysis of subsidy competition in the U.S.. I first showed that states have strong incentives to subsidize firm relocations in order to gain at the expense of other states. I then showed that observed subsidies are much closer to cooperative subsidies than non-cooperative subsidies but that the potential costs of an escalation of subsidy competition are large.

As with all calibration studies, my quantitative results are best interpreted as rough estimates which have to be taken with a grain of salt. The reason is simply that they are obtained from a theoretical model with numbers which abstracts from many features of reality. Having said this, they still provide the best guess available from the academic literature to date of the potential gains and losses from more or less subsidy competition in the U.S.. As such, they hopefully serve as a useful input into policy discussions as well as a useful benchmark for future academic research.

While I used my framework to study subsidy competition among regional governments, it should be clear that it can also be applied to study subsidy competition among national governments. In my view, this would be a valuable contribution to the international subsidy competition/tax competition literature in that it would go beyond the usual analysis of fiscal externalities. In particular, it would make the case that national governments care about attracting multinational firms not only because they expand the national tax base but also because they generate spillover effects for the national economy.

6 Appendix

6.1 Appendix 1: Isomorphism with Armington model

Introducing only the modifications described in subsection 2.4, it should be easy to verify that out of all the conditions in Definition 1 only (7), (8), (13), and (19) change. In particular, the Armington analog to equation (7) is

$$P_j = \left(\sum_{i=1}^{R} (p_{ij})^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}}$$

since the number of firms is now exogenous and normalized to one. Also, the Armington analog to equation (8) is

$$\eta\theta^L \sum_{j=1}^{R} (p_{ij})^{1-\varepsilon} (P_j)^{\varepsilon-1} E_j = \rho_i w_i L_i$$

which simply says that a fraction $\eta\theta^L$ of firm revenues is spent on worker compensation. The Armington analog to equation (13) is

$$p_{ij} = \frac{\left((w_i)^{\theta^L} (r_i)^{\theta^T} \right)^{\eta} (P_i)^{1-\eta} \rho_i \tau_{ij}}{\varphi_i Z_i^{\phi}}$$

which should be intuitive since firms no longer charge markups but productivity is now $\varphi_i Z_i^{\phi}$. Finally, the Armington analog to equation (19) is

$$Z_i = \frac{L_i}{\eta\theta^L} \frac{w_i}{\left((w_i)^{\theta^L} (r_i)^{\theta^T} \right)^{\eta} (P_i)^{1-\eta}}$$

which should make sense since $Z_i = M_i z_i$ in the original model and now M_i is exogenous and normalized to one.

Equations (7), (8), (13), and (19) from the main model can be combined into the two condensed equilibrium conditions

$$P_j = \left(\sum_{i=1}^{R} \frac{L_i}{\eta\theta^L} \frac{w_i}{\left(\left((w_i)^{\theta^L} (r_i)^{\theta^T} \right)^{\eta} (P_i)^{1-\eta} \right)^{\varepsilon}} \left(\frac{\rho_i \tau_{ij}}{\tilde{\varphi}_i} \right)^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}}$$

$$\sum_{j=1}^{R} \left(\frac{\tau_{ij}}{\tilde{\varphi}_i} \right)^{1-\varepsilon} (P_j)^{\varepsilon-1} E_j = (\rho_i)^{\varepsilon} \left(\left((w_i)^{\theta^L} (r_i)^{\theta^T} \right)^{\eta} (P_i)^{1-\eta} \right)^{\varepsilon}$$

where I have replaced the original productivity parameter with a rescaled one satisfying $\tilde{\varphi}_i = \left(\frac{\varepsilon^{\varepsilon} f_i}{(\varepsilon-1)^{(\varepsilon-1)}} \right)^{\frac{1}{1-\varepsilon}} \varphi_i$. Similarly, the abovementioned Armington analogs to equations (7), (8), (13), and (19) can be combined into the two condensed equilibrium conditions

$$P_j = \left(\sum_{i=1}^{R} \left(\frac{L_i}{\eta\theta^L} \right)^{\phi(\varepsilon-1)} \frac{(w_i)^{\phi(\varepsilon-1)}}{\left(\left((w_i)^{\theta^L} (r_i)^{\theta^T} \right)^{\eta} (P_i)^{1-\eta} \right)^{(1+\phi)(\varepsilon-1)}} \left(\frac{\rho_i \tau_{ij}}{\varphi_i} \right)^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}}$$

$$\sum_{j=1}^{R} \left(\frac{\tau_{ij}}{\varphi_i} \right)^{1-\varepsilon} (P_j)^{\varepsilon-1} E_j = \left(\frac{w_i L_i}{\eta\theta^L} \right)^{1-\phi(\varepsilon-1)} (\rho_i)^{\varepsilon} \left(\left(\left((w_i)^{\theta^L} (r_i)^{\theta^T} \right)^{\eta} (P_i)^{1-\eta} \right)^{1+\phi} \right)^{\varepsilon-1}$$

where I have left the original productivity parameter unchanged. The isomorphism can now be seen by imposing $\phi = \frac{1}{\varepsilon-1}$ on the condensed Armington conditions which reveals that both models are identical up to the scale of φ_i.

6.2 Appendix 2: Equilibrium conditions with federal subsidies

As discussed in subsection 3.2, I introduce a federal subsidy $s^I = 1/\varepsilon$ on final and intermediate consumption to correct a markup distortion faced by consumers and firms. With such a subsidy, the equilibrium conditions in levels and changes summarized in Definition 1 and Definition 2 extend to:

Definition 1 (extended) *For given subsidies and a numeraire $i \equiv 1$, an equilibrium in levels is a set of $\{P_i, \lambda_i^L, \lambda_i^K, \lambda_i^C\}$ such that*

$$\lambda_i^L = \frac{U_i^{\sigma}}{\sum_{j=1}^{R} U_j^{\sigma}}$$

$$P_j = \left(\sum_{i=1}^{R} M_i \left(p_{ij}\right)^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}}$$

$$\frac{1}{\varepsilon}\sum_{j=1}^{R}\left(p_{ij}\right)^{1-\varepsilon}\left(P_j\right)^{\varepsilon-1}E_j = \left(\left(w_i\right)^{\theta^L}\left(r_i\right)^{\theta^T}\right)^{\eta}\left(\rho^I P_i\right)^{1-\eta}\rho_i f_i$$

$$r_i T_i = \frac{\mu}{1-\mu}\rho^F E_i^F + \frac{\eta\theta^T}{1-\eta}\rho^F E_i^I$$

where

$$w_i = \frac{\lambda_i^K}{\lambda_i^L}\frac{\theta^L}{\theta^K}\frac{K}{L}$$

$$r_i = \frac{\lambda_i^K}{\lambda_i^C}\frac{\theta^T}{\theta^K}\frac{K}{T_i}$$

$$E_i^I = \frac{1-\eta}{\rho^F\eta\theta^K}\lambda_i^K K$$

$$p_{ij} = \frac{\varepsilon}{\varepsilon-1}\frac{\left(\left(w_i\right)^{\theta^L}\left(r_i\right)^{\theta^T}\right)^{\eta}\left(\rho^F P_i\right)^{1-\eta}\rho_i\tau_{ij}}{\varphi_i}$$

$$S_i = \left(s_i\lambda_i^K + \lambda_i^L s^F\sum_m \rho_m\lambda_m^K\right)\frac{K}{\eta\theta^K}$$

$$\Omega_i = \rho^F N X_i - \left(\left(\lambda_i^K - \lambda_i^L\right)K - \left(\frac{\sum_n X_{in}}{\sum_m\sum_n X_{mn}} - \lambda_i^L\right)s^F\sum_m\sum_n X_{mn}\right)$$

$$\rho^F E_i^F = \left(1-\mu\right)\left(w_i L_i + \lambda_i^L K + r_i T_i - \left(S_i + \Omega_i\right)\right)$$

$$E_i = E_i^F + E_i^I$$

$$U_i = \frac{1}{1-\mu}\frac{A_i}{L_i}\frac{\rho^F E_i^F}{\left(r_i\right)^{\mu}\left(\rho^F P_i\right)^{1-\mu}}$$

$$M_i = \frac{L_i}{\varepsilon f_i\eta\theta^L}\frac{w_i}{\left(\left(w_i\right)^{\theta^L}\left(r_i\right)^{\theta^T}\right)^{\eta}\left(\rho^F P_i\right)^{1-\eta}}$$

Definition 2 (extended) *For given subsidy changes and a numeraire $i \equiv 1$, an equilibrium in changes is a set of $\left\{\hat{P}_i, \hat{\lambda}_i^L, \hat{\lambda}_i^K, \hat{\lambda}_i^C\right\}$ such that*

$$\hat{\lambda}_i^L = \frac{\left(\hat{U}_i\right)^{\sigma}}{\sum_{j=1}^{R}\lambda_j^L\left(\hat{U}_j\right)^{\sigma}}$$

$$\hat{P}_j = \left(\sum_{i=1}^{R} \alpha_{ij} \hat{M}_i \left(\hat{p}_{ij} \right)^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}}$$

$$\sum_{j=1}^{R} \beta_{ij} \left(\hat{p}_{ij} \right)^{1-\varepsilon} \left(\hat{P}_j \right)^{\varepsilon-1} \hat{E}_j = \left(\left(\hat{w}_i \right)^{\theta^L} \left(\hat{r}_i \right)^{\theta^T} \right)^{\eta} \left(\hat{P}_i \right)^{1-\eta} \hat{\rho}_i$$

$$\hat{r}_i = \left(1 - \lambda_i^C \right) \hat{E}_i^F + \lambda_i^C \hat{E}_i^I$$

where

$$\hat{w}_i = \frac{\hat{\lambda}_i^K}{\hat{\lambda}_i^L}$$

$$\hat{r}_i = \frac{\hat{\lambda}_i^K}{\hat{\lambda}_i^C}$$

$$\hat{E}_i^I = \hat{\lambda}_i^K$$

$$\hat{p}_{ij} = \left(\left(\hat{w}_i \right)^{\theta^L} \left(\hat{r}_i \right)^{\theta^T} \right)^{\eta} \left(\hat{P}_i \right)^{1-\eta} \hat{\rho}_i$$

$$S_i' = \left(s_i' \lambda_i^K \hat{\lambda}_i^K + \lambda_i^L \hat{\lambda}_i^L s^F \sum_m \rho_m' \lambda_m^K \hat{\lambda}_m^K \right) \frac{K}{\eta \theta^K}$$

$$\hat{E}_i^F = \frac{1-\mu}{\rho^F} \left(\frac{w_i L_i}{E_i^F} \hat{w}_i \hat{\lambda}_i^L + \lambda_i^L \hat{\lambda}_i^L \frac{iK}{E_i^F} \hat{i} + \frac{r_i T_i}{E_i^F} \hat{r}_i - \frac{S_i' + \Omega_i'}{E_i^F} \right)$$

$$\hat{E}_i = \frac{E_i^F}{E_i} \hat{E}_i^F + \frac{E_i^I}{E_i} \hat{E}_i^I$$

$$\hat{U}_i = \frac{1}{\hat{\lambda}_i^L} \frac{\hat{E}_i^F}{\left(\hat{r}_i \right)^{\mu} \left(\hat{P}_i \right)^{1-\mu}}$$

$$\hat{M}_i = \frac{\hat{w}_i \hat{\lambda}_i^L}{\left(\left(\hat{w}_i \right)^{\theta^L} \left(\hat{r}_i \right)^{\theta^T} \right)^{\eta} \left(\hat{P}_i \right)^{1-\eta}}$$

and

$$\alpha_{ij} = \frac{X_{ij}}{\sum_{m=1}^{R} X_{mj}}$$

$$\beta_{ij} = \frac{X_{ij}}{\sum_{n=1}^{R} X_{in}}$$

$$w_i L_i = \frac{\eta \theta^L}{\rho_i} \sum_{n=1}^{R} X_{in}$$

$$K_i = \frac{\eta\theta^K}{\rho_i}\sum_{n=1}^{R} X_{in}$$

$$r_i T_i^C = \frac{\eta\theta^T}{\rho_i}\sum_{n} X_{in}$$

$$E_i^I = \frac{1-\eta}{\rho_i \rho^F}\sum_{n=1}^{R} X_{in}$$

$$E_i^F = \sum_{m=1}^{R} X_{mi} - E_i^I$$

$$E_i = E_i^F + E_i^I$$

$$r_i T_i = \frac{\mu}{1-\mu}\rho^F E_i^F + r_i T_i^C$$

$$\lambda_i^K = \frac{K_i}{\sum_{i=1}^{R} K_i}$$

$$\lambda_i^C = \frac{r_i T_i^C}{r_i T_i}$$

6.3 Appendix 3: Model fit

Appendix Figure 1 illustrates that the variation in trade flows is largely explained by variation in trade costs by plotting the (log) export shares from Illinois against the (log) trade costs from Illinois. The trade costs are backed out using the Head-Ries index $\tau_{ij} = \left(\frac{X_{ij}X_{ji}}{X_{ii}X_{jj}}\right)^{\frac{1}{2(1-\varepsilon)}}$, which follows from equations $X_{ij} = M_i \left(p_{ij}\right)^{1-\varepsilon} \left(P_j\right)^{\varepsilon-1} E_j$ and (13) under the assumption that $\tau_{ij} = \tau_{ji}$ and $\tau_{ii} = 1$. Appendix Figure 2 then shows that these trade costs are highly correlated with distance, just as one would expect.

Appendix Figure 3 illustrates that variation in manufacturing employment is largely explained by variation in amenities with Wyoming and California having the worst and best amenities, respectively. Relative amenities are backed out using the formula $\frac{A_i}{A_j} = \left(\frac{\lambda_i^L}{\lambda_j^L}\right)^{\frac{1+\sigma}{\sigma}} \frac{\left(\frac{r_i}{r_j}\right)^{\mu}\left(\frac{P_i}{P_j}\right)^{1-\mu}}{\frac{E_i^F}{E_j^F}}$ which follows from equations (6) and (18). $\frac{\lambda_i^L}{\lambda_j^L}$ and $\frac{E_i^F}{E_j^F}$ can be directly read off of the data keeping in mind that $E_i^F = \sum_{m=1}^{R} X_{mi} - E_i^I$ and $E_i^I = \frac{1-\eta}{\rho_i\rho^F}\sum_{n=1}^{R} X_{in}$. $\frac{P_i}{P_j}$ can be calculated from $\frac{P_i}{P_j} = \left(\sum_m \alpha_{mj}\left(\frac{\tau_{mi}}{\tau_{mj}}\right)^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}}$ using the trade costs from the

Head-Ries index which follows straightforwardly from equation (7). $\frac{r_i}{r_j}$ is calculated from $\frac{r_i}{r_j} = \frac{\frac{\mu}{1-\mu}E_i^F + \frac{\eta\theta^T}{1-\eta}E_i^I}{\frac{\mu}{1-\mu}E_j^F + \frac{\eta\theta^T}{1-\eta}E_j^I}\frac{T_j}{T_i}$ using state land areas as proxies for T_i.

6.4 Appendix 4: Decomposition of welfare effects

Differentiating equation (18) yields:

$$\frac{dU_j}{U_j} = \frac{dE_j^F}{E_j^F} - \frac{d\lambda_j^L}{\lambda_j^L} - \frac{dP_j}{P_j} - \mu\left(\frac{dr_j}{r_j} - \frac{dP_j}{P_j}\right)$$

Equations (13) and (8) imply $\sum_{n=1}^{R} X_{jn} = M_j(\varepsilon - 1)p_{jj}\varphi_j f_j$. Since $E_j = \sum_{m=1}^{R} X_{mj}$, one can write $E_j = M_j(\varepsilon - 1)p_{jj}\varphi_j f_j - NX_j$. Around $NX_j = 0$, one therefore obtains:

$$\frac{dE_j}{E_j} = \frac{dM_j}{M_j} + \frac{dp_{jj}}{p_{jj}} - \frac{dNX_j}{E_j}$$

Recall from the discussion of equation (19) in the main text that the number of firms can be expressed as $M_j = \frac{1}{\varepsilon f_j}\left(\frac{1}{\eta}\left(\frac{L_j}{\theta^L}\right)^{\theta^L}\left(\frac{K_j}{\theta^K}\right)^{\theta^K}\left(\frac{T_j^C}{\theta^T}\right)^{\theta^T}\right)^{\eta}\left(\frac{C_j^I}{1-\eta}\right)^{1-\eta}$. Exploiting the fact that $C_j^I = \frac{E_I^I}{P_j}$, this implies:

$$\frac{dM_j}{M_j} = \eta\left(\theta^L\frac{d\lambda_j^L}{\lambda_j^L} + \theta^K\frac{d\lambda_j^K}{\lambda_j^K} + \theta^T\frac{d\lambda_j^C}{\lambda_j^C}\right) + (1-\eta)\left(\frac{dE_j^I}{E_j^I} - \frac{dP_j}{P_j}\right)$$

Differentiating equation (7) yields:

$$\frac{dP_j}{P_j} = \sum_{i=1}^{R}\alpha_{ij}\left(\frac{dp_{ii}}{p_{ii}} - \frac{1}{\varepsilon - 1}\frac{dM_i}{M_i}\right)$$

These four equations can be combined to:

$$\begin{aligned}
\frac{dU_j}{U_j} &= \frac{1}{\eta}\sum_{i=1}^{R}\alpha_{ij}\frac{1}{\varepsilon-1}\frac{dM_i}{M_i}+\frac{1}{\eta}\sum_{i=1}^{R}\alpha_{ij}\left(\frac{dp_{jj}}{p_{jj}}-\frac{dp_{ii}}{p_{ii}}\right)\\
&\quad -\mu\left(\frac{dr_j}{r_j}-\frac{dP_j}{P_j}\right)-\theta^T\left(\frac{d\lambda_j^L}{\lambda_j^L}-\frac{d\lambda_j^C}{\lambda_j^C}\right)\\
&\quad +\frac{1}{\eta}\left((1-\eta)\frac{dE_j^I}{E_j^I}+\eta\frac{dE_j^F}{E_j^F}-\frac{dE_j}{E_j}\right)\\
&\quad +\frac{1}{\eta}\left(\eta\theta^K\left(\frac{d\lambda_j^K}{\lambda_j^K}-\frac{d\lambda_j^L}{\lambda_j^L}\right)-\frac{dNX_j}{E_j}\right)
\end{aligned}$$

This equation simplifies to equation (44) in the main text because the last two terms are equal to zero around $s_i = \Omega_i = NX_i = 0$. To see why the second to last term is zero, notice that $E_i^F/E_i = \eta$ and $E_i^I/E_i = 1-\eta$ if $s_i = \Omega_i = NX_i = 0$ since then $E_i^F = \frac{K_i}{\theta^K}$, $E_i^I = \frac{1-\eta}{\eta}\frac{K_i}{\theta^K}$, and $E_i = \frac{K_i}{\eta\theta^K}$ as follows from combining equations (9), (10), (11), (15), (16), and (17). To see why the last term is zero, make use of some of the same relationships, namely (15) and $E_i = \frac{K_i}{\eta\theta^K}$.

6.5 Appendix 5: Algorithm

I compute the optimal subsidies of state i by solving $\min_{\left\{s_i', \hat{P}_j, \hat{\lambda}_j^L, \hat{\lambda}_j^K, \hat{\lambda}_j^C\right\}_{j=1,\ldots,R}} -\hat{U}_i$ subject to the equilibrium conditions in changes as summarized in Definition 2 (extended) in Appendix 2. Notice that minimizing $-\hat{U}_i$ is equivalent to maximizing U_i which is, in turn, equivalent to maximizing G_i^{loc} from equation (3). This follows the approach of Su and Judd (2012) which builds on the idea of mathematical programming with equilibrium constraints.

I compute Nash subsidies following the same method I applied in Ossa (2014). Starting at factual subsidies, I compute each state's optimal subsidies, then impose these optimal subsidies, and let all states reoptimize given all other states' optimal subsidies, and so on, until the solution converges in the sense that no state has an incentive to deviate from its subsidies. I have experimented with many different starting values without finding any differences in the results which makes me believe that the identified Nash equilibrium is unique.

I compute cooperative transfers and subsidies analogously to optimal subsidies by solving $\min_{\left\{s_i',\Omega_i',\hat{P}_i,\hat{\lambda}_i^L,\hat{\lambda}_i^K,\hat{\lambda}_i^C\right\}_{i=1,\ldots,R}} -\left(\sum_{j=1}^{R}\lambda_j^L\left(\hat{U}_j\right)^{\sigma}\right)^{\frac{1}{\sigma}}$ subject to the equilibrium conditions in changes as summarized in Definition 2 (extended) in Appendix 2. Notice that minimizing $-\left(\sum_{j=1}^{R}\lambda_j^L\left(\hat{U}_j\right)^{\sigma}\right)^{\frac{1}{\sigma}}$ is equivalent to maximizing $\left(\sum_{j=1}^{R}U_j^{\sigma}\right)^{\frac{1}{\sigma}}$ which is, in turn, equivalent to maximizing G^{fed} from equation (3). To accelerate convergence, I provide analytic derivatives of the objective functions and the equilibrium constraints throughout.

References

[1] Allen, Treb and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *Quarterly Journal of Economics* 129(3): 1085-1140.

[2] Armington, Paul. 1969. "A Theory of Demand for Products Distinguished by Place of Production." *IMF Staff Papers* 16: 159-176.

[3] Baldwin, Richard, Rikard Forslid, Philippe Martin, Gianmarco Ottaviano, and Frederic Robert-Nicoud. 2005. *Economic Geography and Public Policy.* Princeton, NJ: Princeton University Press.

[4] Bartik, Timothy. 2017. "A New Panel Database on Business Incentives for Economic Development Offered by State and Local Governments in the United States." Prepared for the Pew Charitable Trusts.

[5] Caliendo, Lorenzo, Fernando Parro, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte. 2014. "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy." NBER Working Paper No. 20168.

[6] Davis, Morris and Francois Ortalo-Magne. 2011. "Household Expenditures, Wages, Rents." *Review of Economic Dynamics* 14: 248-261.

[7] Dekle, Robert, Jonathan Eaton, and Sam Kortum. 2007. "Unbalanced Trade." *American Economic Review Papers and Proceedings* 97(2): 351-355.

[8] Fajgelbaum, Pablo, Eduardo Morales, Juan Carlos Suarez Serrato, and Owen Zidar. 2016. "State Taxes and Spatial Misallocation." *NBER Working Paper 21760.*

[9] Fujita, Masahisa, Paul Krugman, and Anthony Venables. 2001. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.

[10] Gaubert, Cecile. 2014. "Firm Sorting and Agglomeration." University of California at Berkeley, Mimeo.

[11] Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings." *Journal of Political Economy* 118(3): 536-598.

[12] Handbury, Jesse and David Weinstein. 2015. "Goods Prices and Availability in Cities." *The Review of Economic Studies* 82(1): 258-296.

[13] Kline, Patrick and Enrico Moretti. 2014. "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *Quarterly Journal of Economics* 129(1): 275-331.

[14] Krugman, Paul. 1991. "Increasing Returns and Economic Geography." *Journal of Political Economy* 99(3): 438-499.

[15] Krugman, Paul and Anthony Venables. 1995. "Globalization and the Inequality of Nations." *Quarterly Journal of Economics* 110(4): 857-880.

[16] Oates, Wallace. 1972. *Fiscal Federalism*. Hartcourt Brace Jovanovich, New York.

[17] Ossa, Ralph. 2011. "A "New Trade" Theory of GATT/WTO Negotiations." *Journal of Political Economy* 119(1): 122-152.

[18] Ossa, Ralph. 2012. "Profits in the New Trade Approach to Trade Negotiations." *American Economic Review Papers and Proceedings* 104(12): 4104-4146.

[19] Ossa, Ralph. 2014. "Trade Wars and Trade Talks with Data." *American Economic Review* 104(12): 4104-46.

[20] Redding, Stephen. 2016. "Goods Trade, Factor Mobility and Welfare." *Journal of International Economics* 101: 148-167.

[21] Rogers, James. 2000. "The Effectiveness and Constitutionality of State Tax Incentive Policies for Locating Businesses: A Simple Game Theoretic Analysis." *The Tax Lawyer* 53(2): 431-458.

[22] Suarez Serrato, Juan and Owen Zidar. 2016, "Who Benefits from Corporate Tax Cuts? A Local Labor Market Approach with Heterogeneous Firms." *American Economic Review* 106(9): 2582-2624.

[23] Story, Louise, Tiff Fehr, and Derek Watkins. 2012. "Explore Government Subsidies." *The New York Times.* Accessed using: http://www.nytimes.com/interactive/2012/12/01/us/government-incentives.html.

[24] Su, Che-Lin and Kenneth L. Judd. 2012. "Constrained Optimization Approaches to Estimation of Structural Models." *Econometrica* 80(5): 2213-2230.

[25] Thomas, Kenneth. 2011. "Investment Incentives and the Global Competition for Capital. *New York: Palgrove-MacMillan.*

[26] Venables, Anthony. 1987. "Trade and Trade Policy with Differentiated Products: A Chamberlinian-Ricardian Model." *The Economic Journal* 97(387): 700-717.

Figure 1: Grid search for multiple equilibria

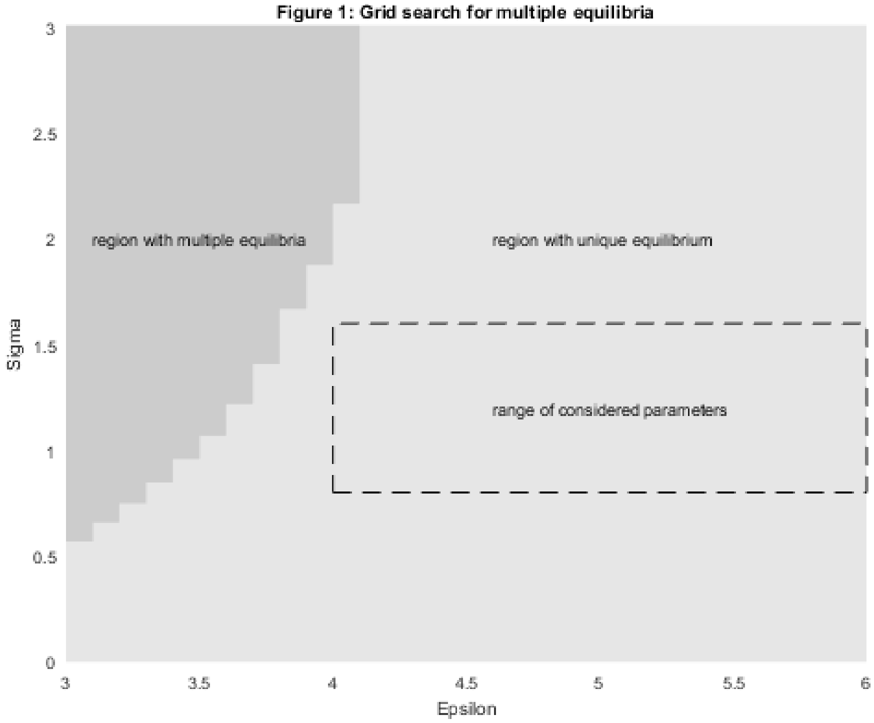


Figure 2: Effects of subsidy imposed by IL

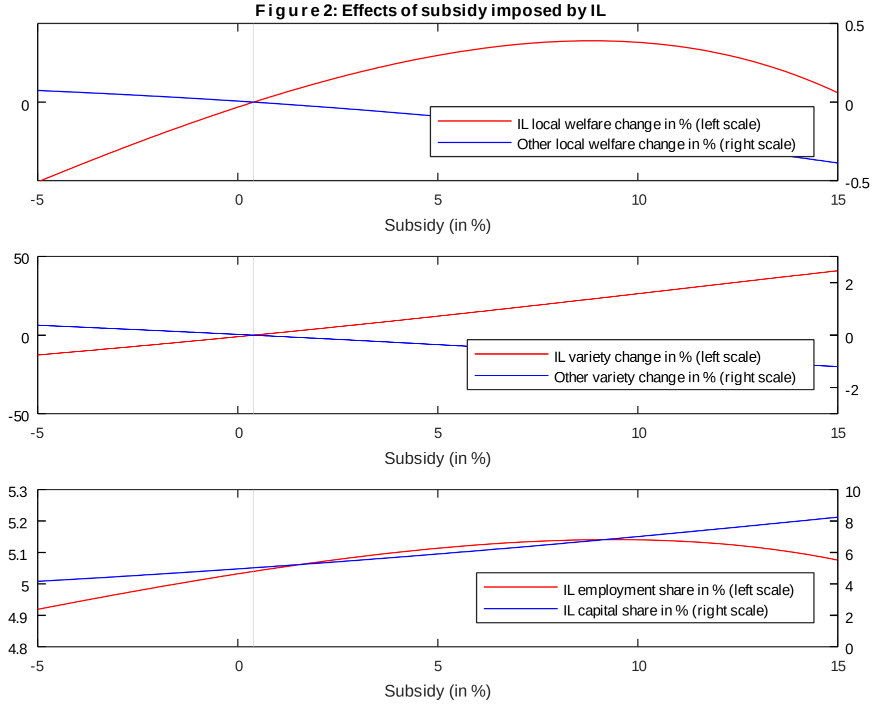


Figure 3: Optimal subsidies

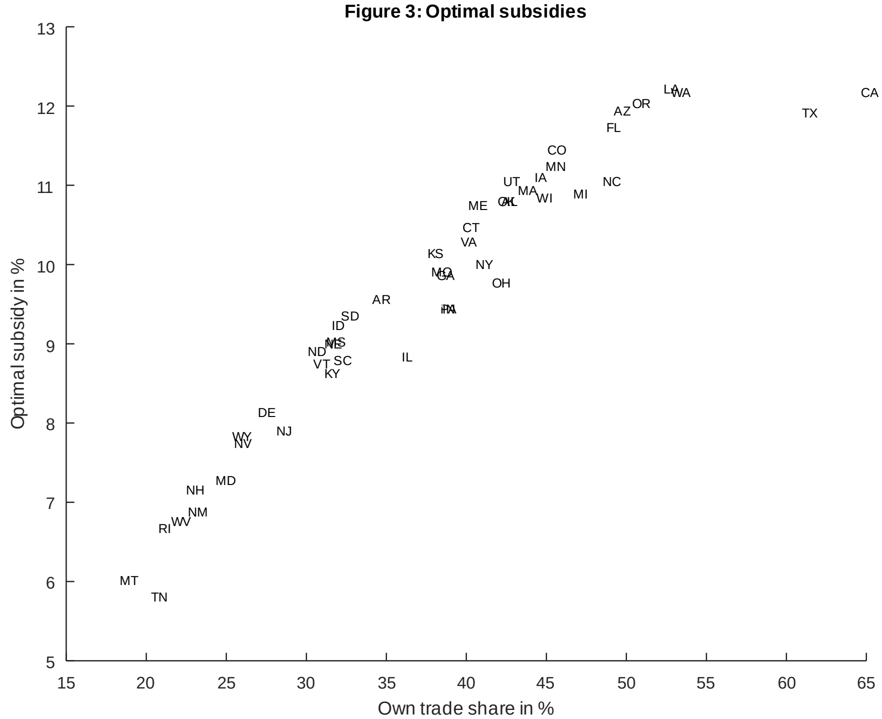


Figure 4: Welfare gains of optimal subsidies

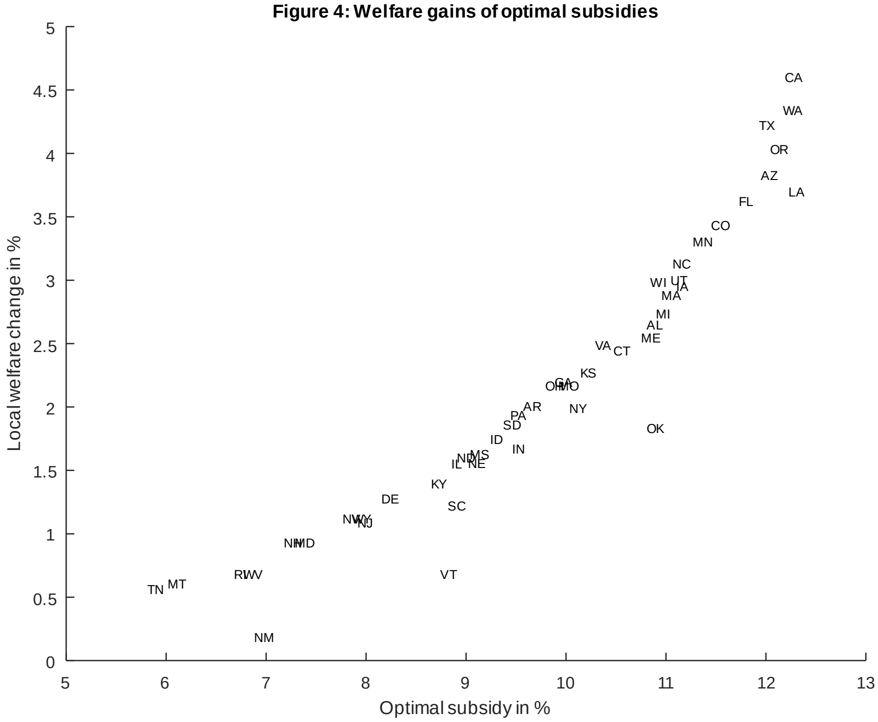


Figure 5: Maximizing employment instead of welfare

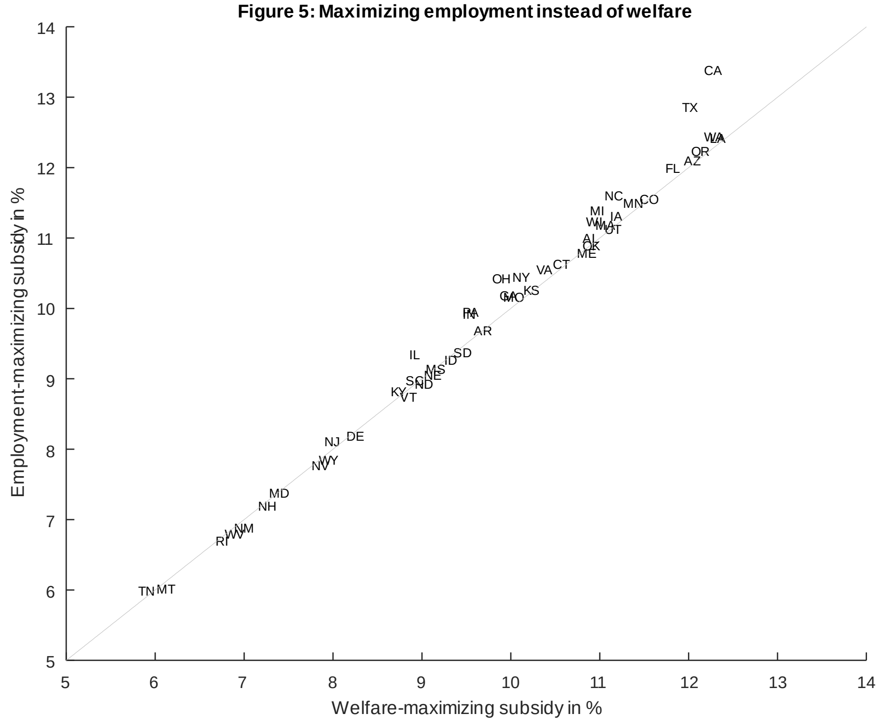


Figure 6: Welfare effects resulting from optimal subsidy imposed by IL

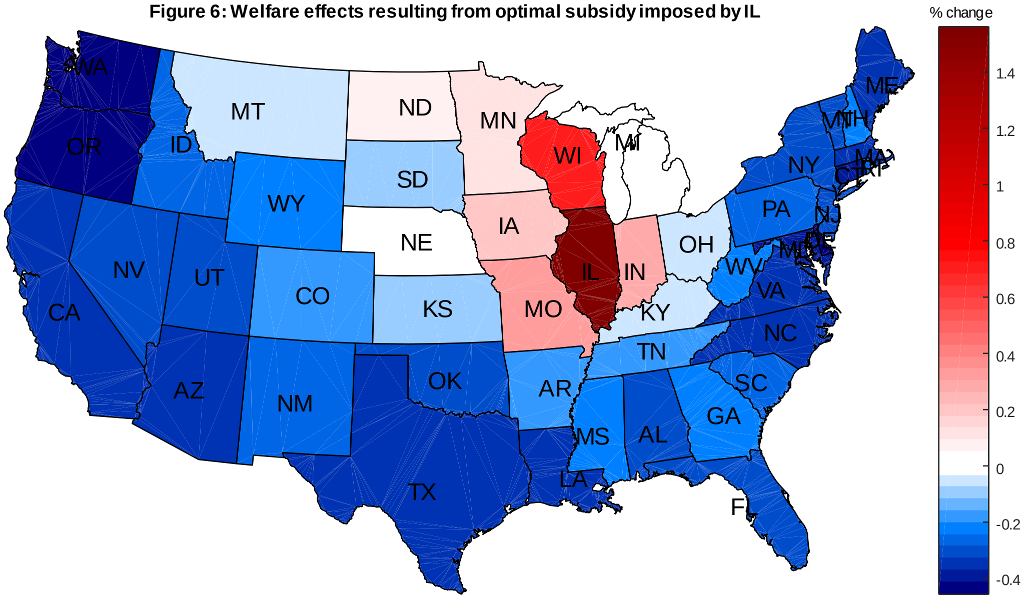


Figure 7: Optimal subsidies w/ and w/o federal subsidies

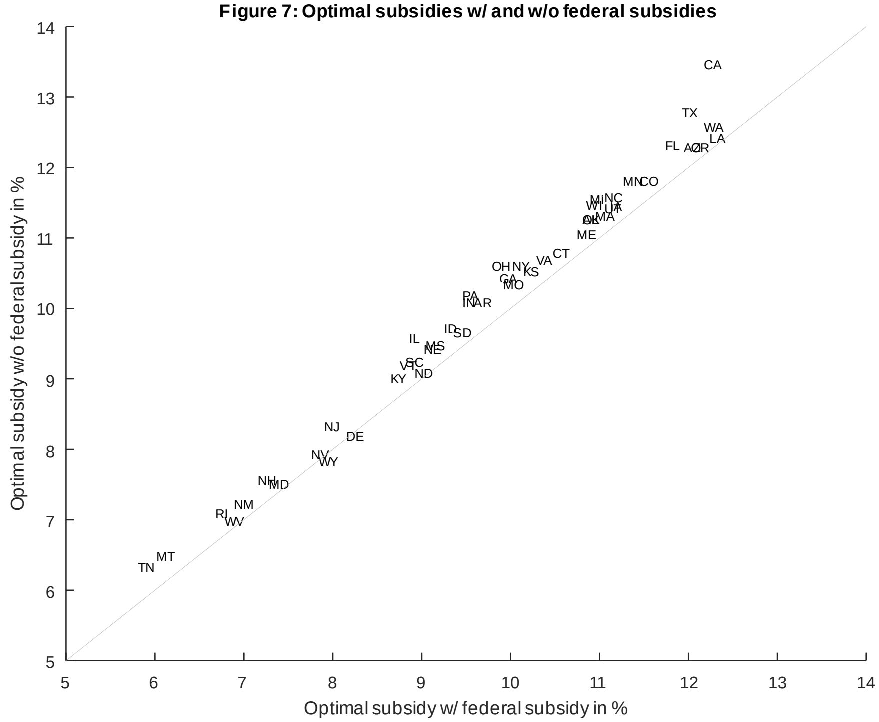


Figure 8: Own welfare gains w/ and w/o federal subsidies

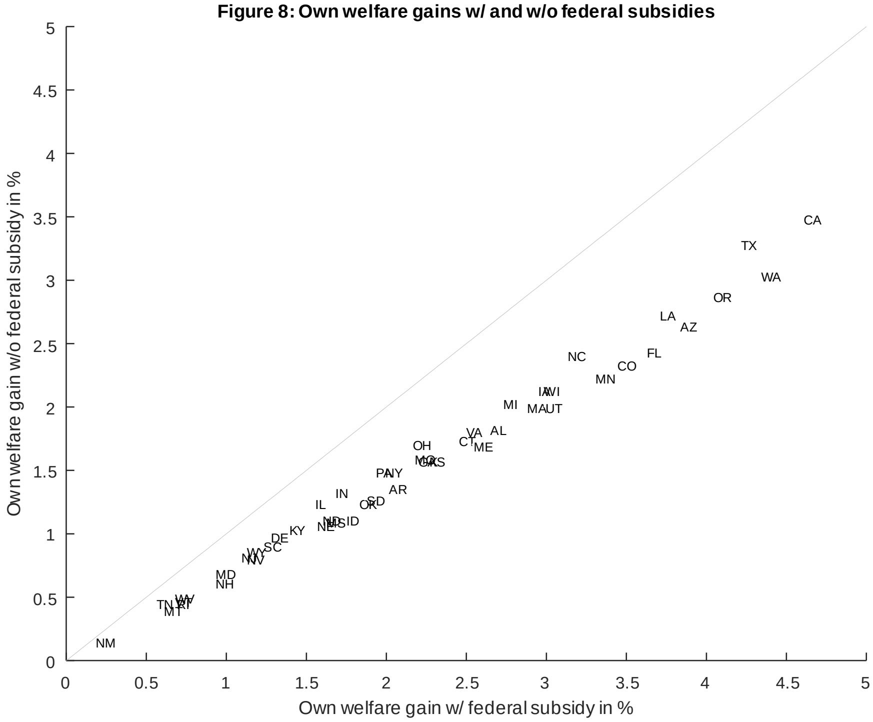


Figure 9: Nash subsidies vs. optimal subsidies

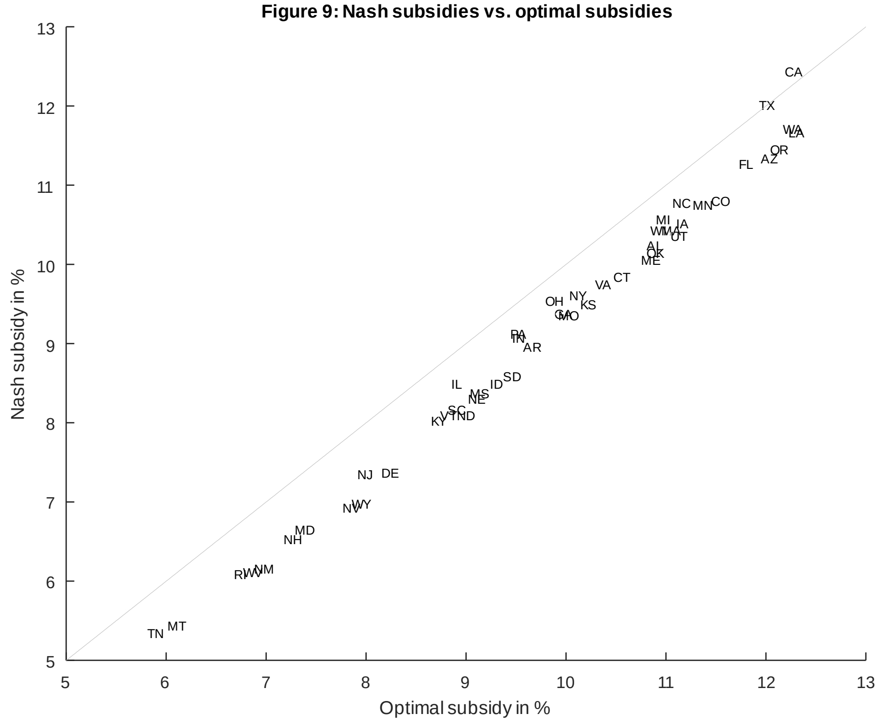


Figure 10: Determinants of welfare change

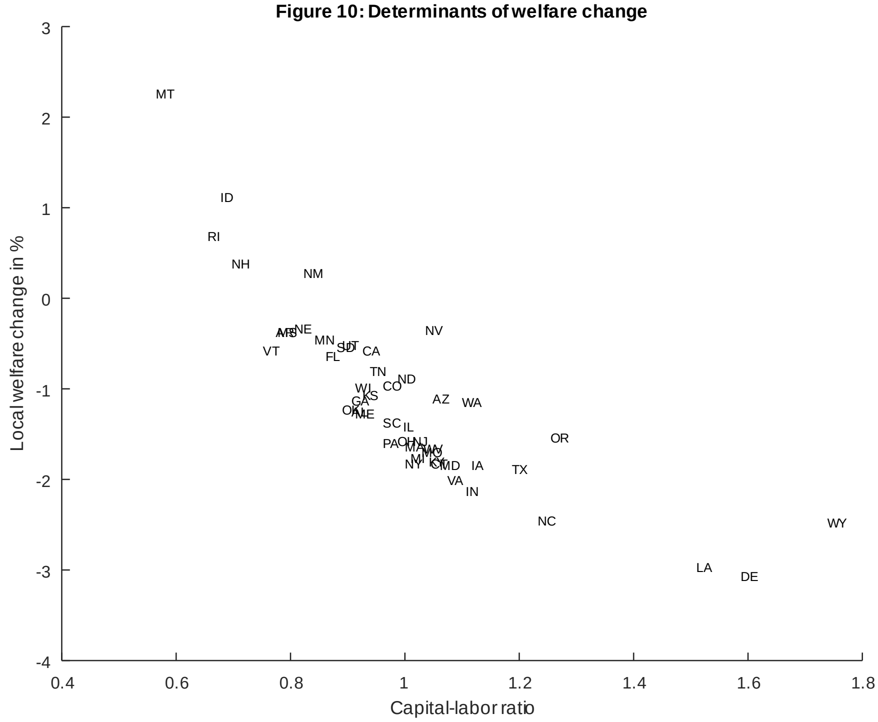


Figure 11: Welfare effects of Nash subsidies

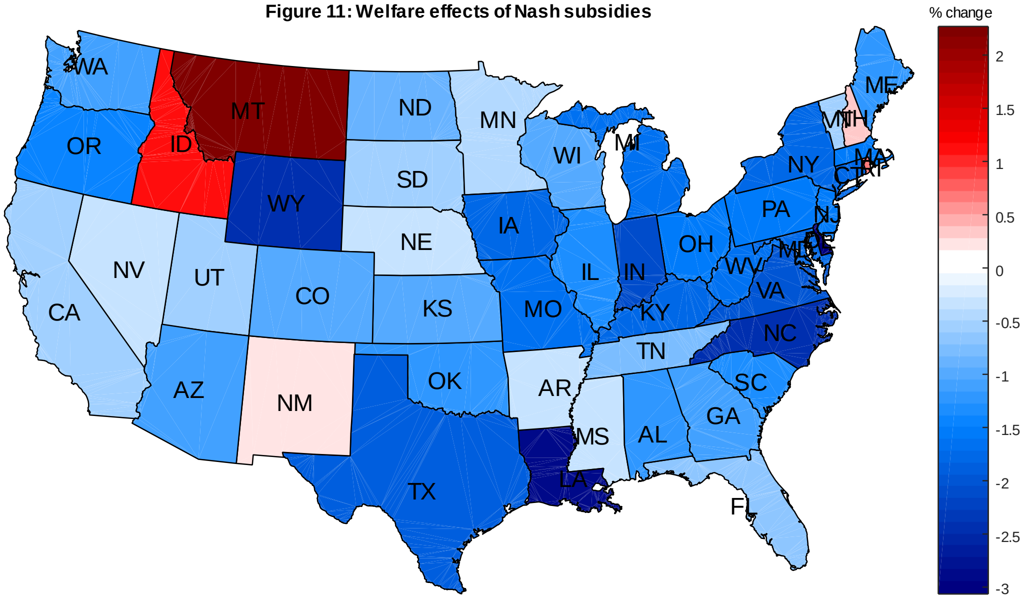


Figure 12: Nash subsidies w/ and w/o federal subsidies

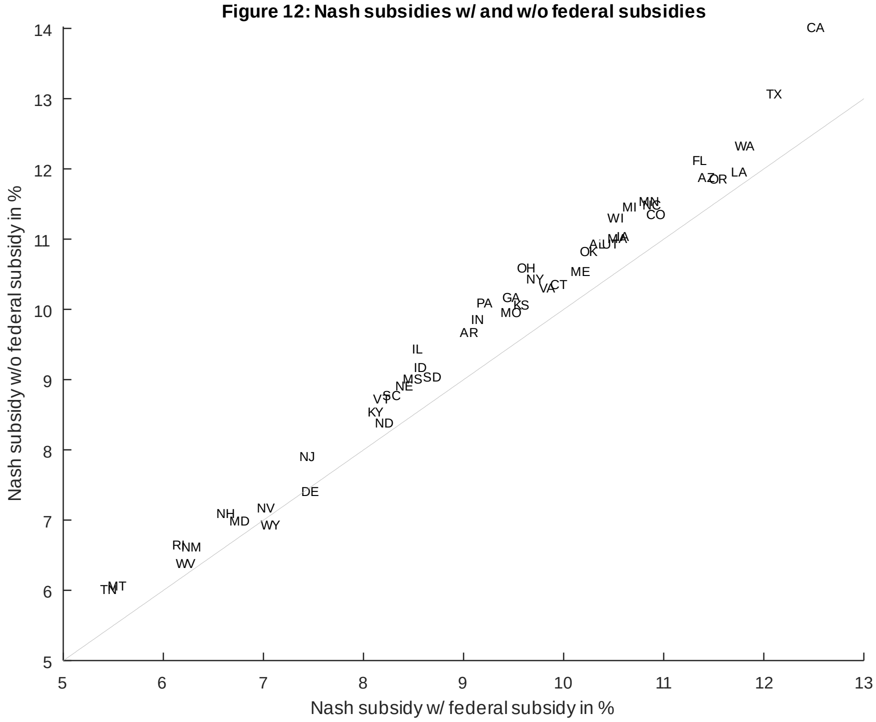


Figure 13: Welfare change w/ and w/o federal subsidies

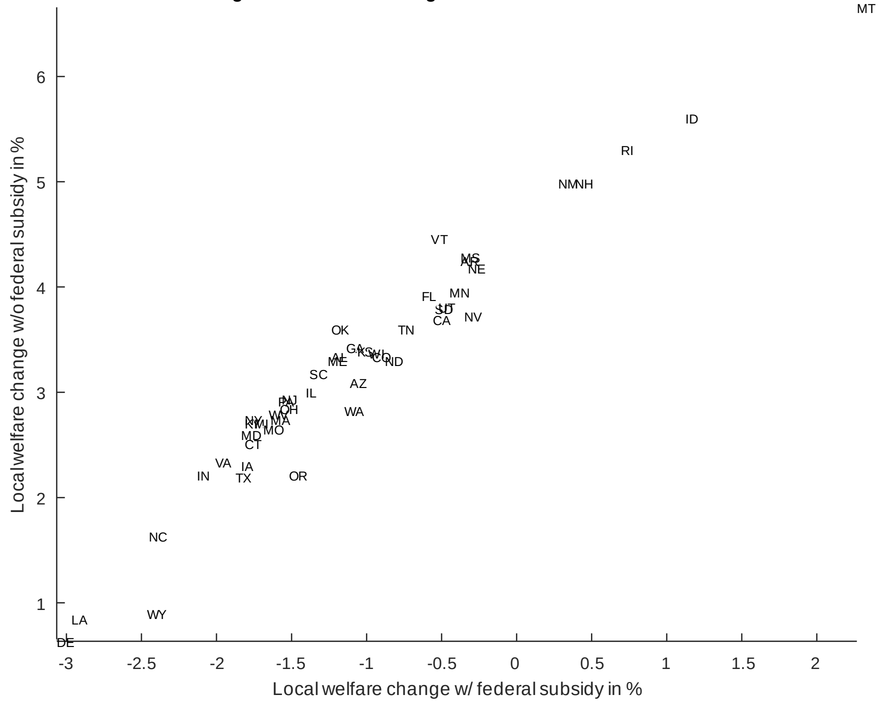


Figure 14: Cooperative redistribution

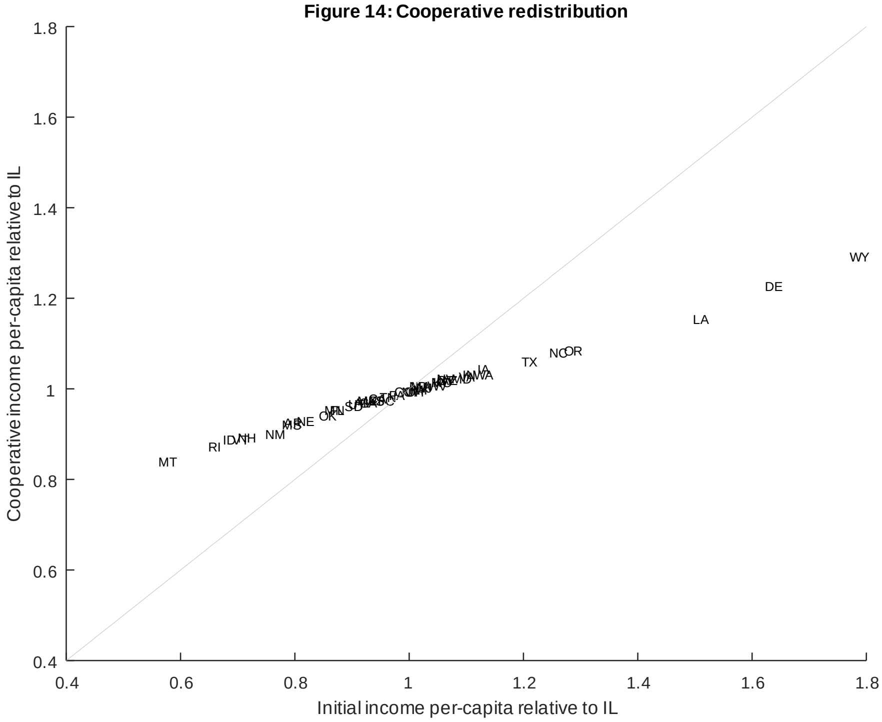


Figure 15: Cooperative subsidies, Nash subsidies, and factual subsidies

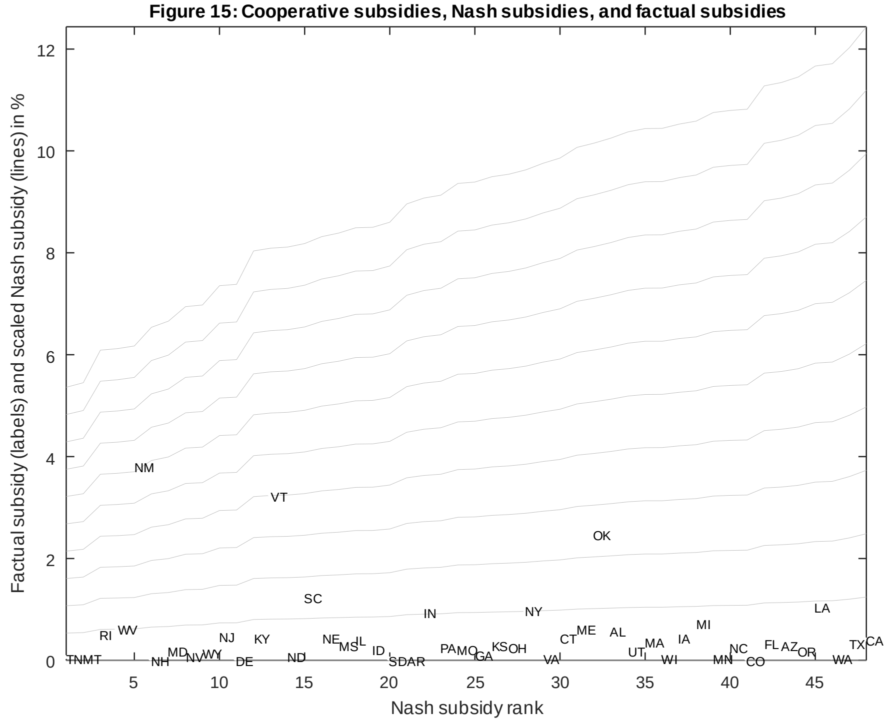


Figure 16: Factual subsidy costs vs. Nash subsidy costs

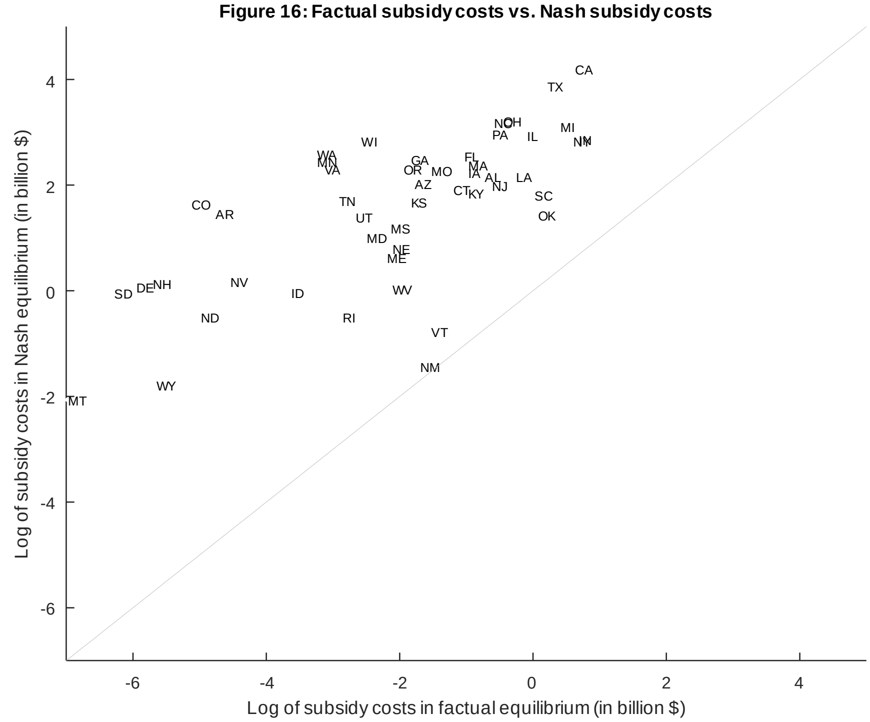
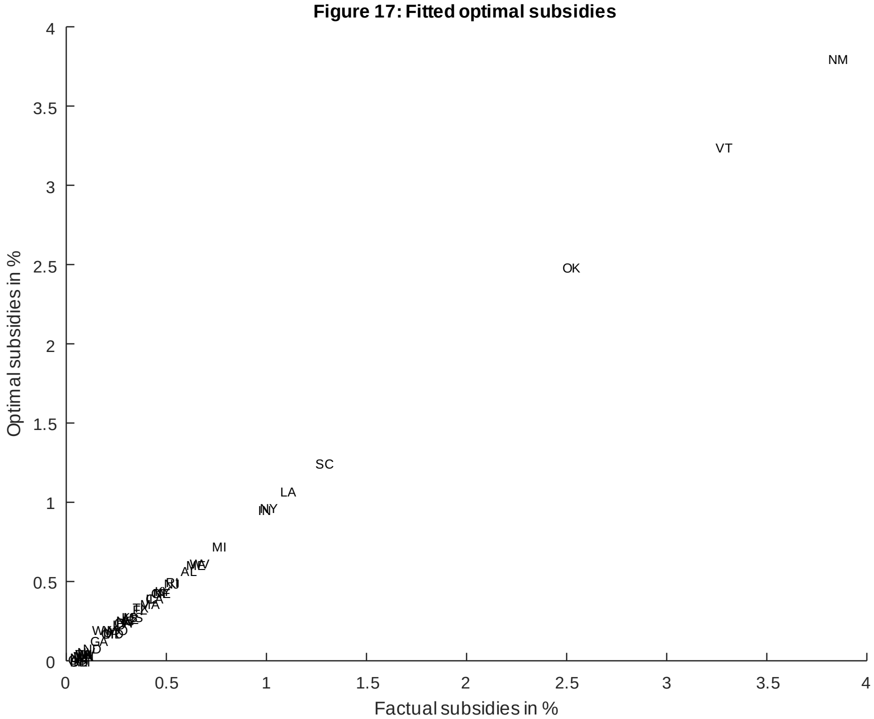
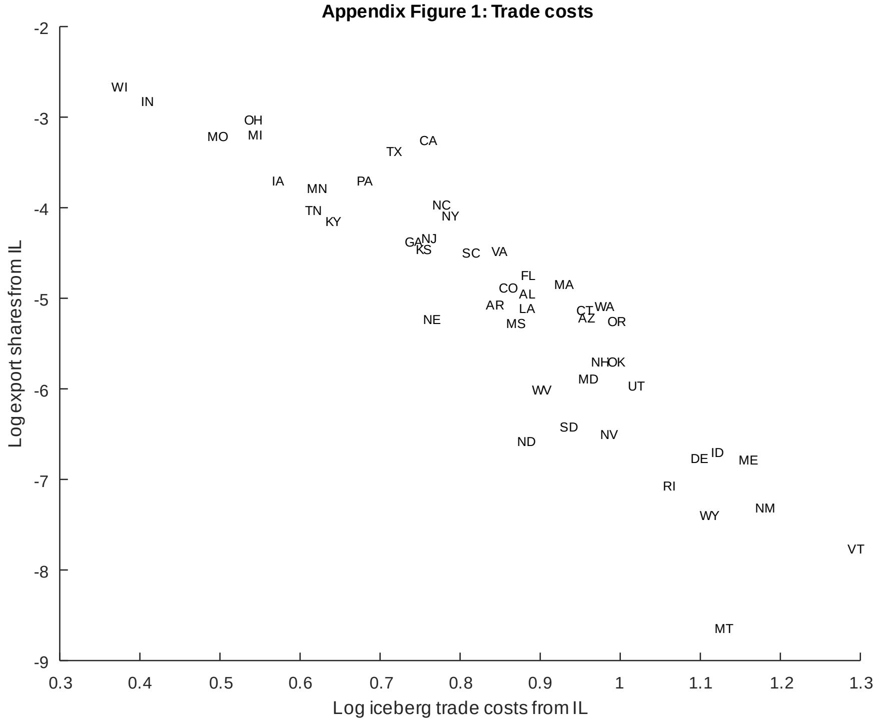


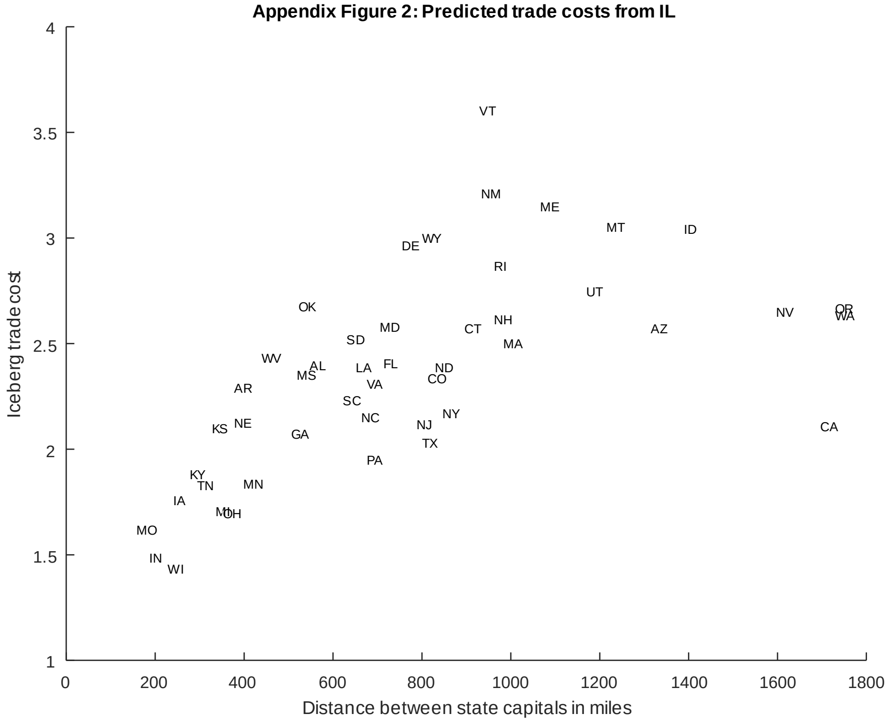
Figure 17: Fitted optimal subsidies



Appendix Figure 1: Trade costs



Appendix Figure 2: Predicted trade costs from IL



Appendix Figure 3: Relative amenities

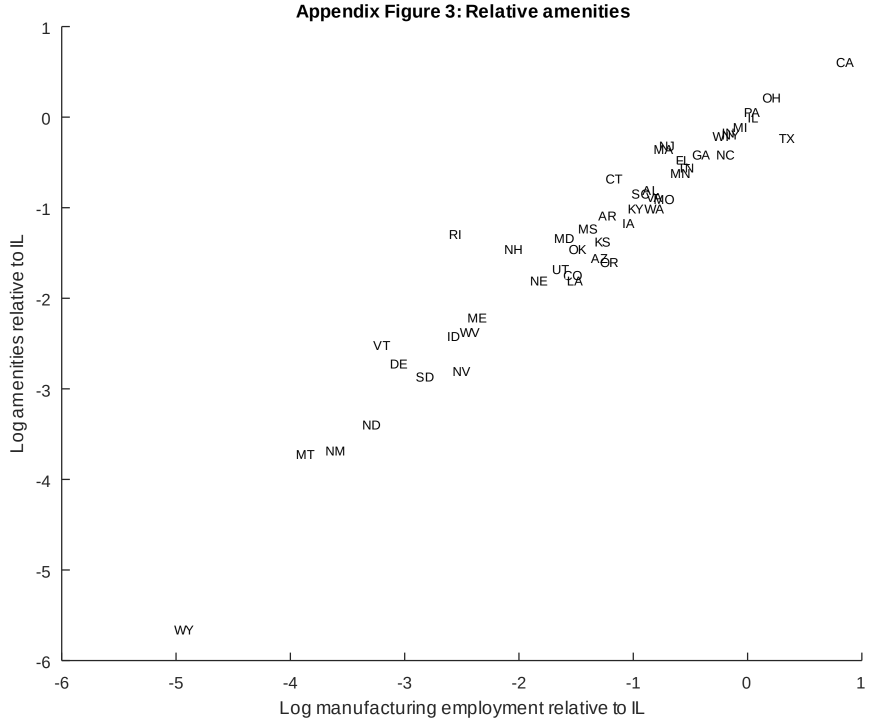


Table 1: Optimal subsidies

	optimal subsidy		Δ welfare				
	subsidy (%)	subsidy ($bn)	own (%)	own ($bn)	other (%)	other ($bn)	national (%)
AL	10.8	13.7	2.6	1.1	-0.16	-2.4	-0.05
AZ	11.9	12.1	3.8	1.1	-0.15	-2.7	-0.07
AR	9.6	6.8	2.0	0.5	-0.08	-1.0	-0.02
CA	12.2	91.8	4.6	10.2	-1.18	-21.9	-0.47
CO	11.5	8.2	3.4	0.7	-0.09	-1.7	-0.04
CT	10.5	10.8	2.4	0.8	-0.15	-1.9	-0.05
DE	8.2	1.8	1.3	0.1	-0.03	-0.3	-0.01
FL	11.7	20.0	3.6	1.8	-0.22	-4.1	-0.09
GA	9.9	18.6	2.2	1.4	-0.23	-3.1	-0.07
ID	9.2	1.5	1.7	0.1	-0.02	-0.2	0.00
IL	8.9	27.9	1.6	1.7	-0.28	-4.0	-0.10
IN	9.5	26.8	1.7	1.6	-0.29	-4.1	-0.12
IA	11.1	14.8	3.0	1.2	-0.16	-2.9	-0.08
KS	10.1	8.4	2.3	0.6	-0.10	-1.4	-0.03
KY	8.6	9.9	1.4	0.6	-0.11	-1.4	-0.04
LA	12.2	14.2	3.7	1.3	-0.20	-3.3	-0.11
ME	10.8	3.0	2.5	0.2	-0.04	-0.5	-0.01
MD	7.3	4.3	0.9	0.2	-0.05	-0.5	-0.01
MA	11.0	17.0	2.9	1.4	-0.21	-3.2	-0.07
MI	10.9	33.5	2.7	2.7	-0.39	-6.3	-0.16
MN	11.3	17.5	3.3	1.6	-0.20	-3.4	-0.07
MS	9.0	5.2	1.6	0.3	-0.07	-0.7	-0.01
MO	9.9	15.2	2.2	1.1	-0.17	-2.6	-0.06
MT	6.0	0.2	0.6	0.0	0.00	0.0	0.00
NE	9.0	3.5	1.6	0.2	-0.04	-0.5	-0.01
NV	7.8	1.9	1.1	0.1	-0.03	-0.2	0.00
NH	7.2	1.8	0.9	0.1	-0.02	-0.2	0.00
NJ	7.9	11.3	1.1	0.6	-0.13	-1.4	-0.04
NM	6.9	0.4	0.2	0.0	0.00	0.0	0.00
NY	10.0	26.2	2.0	1.7	-0.29	-4.2	-0.11
NC	11.1	37.0	3.1	3.3	-0.47	-7.7	-0.23
ND	8.9	1.0	1.6	0.1	-0.01	-0.1	0.00
OH	9.8	36.7	2.2	2.7	-0.42	-6.3	-0.16
OK	10.8	6.5	1.8	0.4	-0.07	-1.0	-0.03
OR	12.0	15.9	4.0	1.6	-0.24	-3.7	-0.11
PA	9.5	28.9	1.9	2.0	-0.31	-4.6	-0.12
RI	6.7	0.9	0.7	0.0	-0.01	-0.1	0.00
SC	8.8	9.5	1.2	0.5	-0.10	-1.3	-0.04
SD	9.4	1.5	1.9	0.1	-0.02	-0.2	0.00
TN	5.8	8.4	0.6	0.3	-0.07	-0.7	-0.01
TX	11.9	69.8	4.2	7.4	-0.85	-16.9	-0.49
UT	11.1	6.5	3.0	0.5	-0.08	-1.2	-0.02
VT	8.7	0.7	0.7	0.0	-0.01	-0.1	0.00
VA	10.3	15.7	2.5	1.2	-0.20	-2.9	-0.08
WA	12.2	20.9	4.3	2.1	-0.29	-5.0	-0.14
WV	6.8	1.6	0.7	0.1	-0.02	-0.2	-0.01
WI	10.9	26.0	3.0	2.2	-0.30	-5.0	-0.11
WY	7.8	0.3	1.1	0.0	0.00	0.0	0.00
Average	9.6	14.9	2.2	1.2	-0.18	-2.9	-0.07

Table 2: Sensitivity checks for optimal subsidies

Panel A: Sensitivity wrt. sigma

	subsidy		Δ welfare	
σ	avg	own	other	national
0.80	9.6	2.2	-0.2	-0.1
1.00	9.6	2.2	-0.2	-0.1
1.20	9.6	2.2	-0.2	-0.1
1.40	9.7	2.1	-0.2	-0.1
1.60	9.7	2.1	-0.2	-0.1

Panel B: Sensitivity wrt. epsilon

	subsidy		Δ welfare	
ε	avg	own	other	national
4.00	13.0	6.7	-0.7	-0.3
4.50	11.0	3.5	-0.3	-0.1
5.00	9.6	2.2	-0.2	-0.1
5.50	8.6	1.5	-0.1	0.0
6.00	7.8	1.1	-0.1	0.0

Panel C: Sensitivity wrt. phi

	subsidy		Δ welfare	
ϕ	avg	own	other	national
0.33	16.4	15.7	-1.5	-0.6
0.29	12.5	5.0	-0.4	-0.2
0.25	9.6	2.2	-0.2	-0.1
0.22	7.4	1.0	-0.1	0.0
0.20	5.6	0.5	0.0	0.0

Panel D: Sensitivity wrt. intial subsidies

state	subsidy		state	subsidy	
	min	max		min	max
AL	10.6	10.8	NE	8.7	9.1
AZ	11.7	12.0	NV	7.4	7.8
AR	9.3	9.6	NH	6.9	7.2
CA	12.2	12.3	NJ	7.7	8
CO	11.2	11.5	NM	6.9	7.2
CT	10.2	10.5	NY	9.9	10.1
DE	7.8	8.2	NC	10.9	11.1
FL	11.5	11.8	ND	8.6	8.9
GA	9.6	9.9	OH	9.6	9.8
ID	8.9	9.3	OK	10.7	11
IL	8.7	8.9	OR	11.8	12
IN	9.3	9.5	PA	9.3	9.5
IA	10.9	11.1	RI	6.4	6.7
KS	9.9	10.2	SC	8.6	8.9
KY	8.4	8.7	SD	9	9.4
LA	12.1	12.3	TN	5.6	5.8
ME	10.5	10.8	TX	11.9	12
MD	7.0	7.3	UT	10.8	11.1
MA	10.7	11.0	VT	8.7	9
MI	10.8	10.9	VA	10	10.3
MN	11.0	11.3	WA	12	12.2
MS	8.7	9.1	WV	6.5	6.8
MO	9.7	9.9	WI	10.6	10.9
MT	5.7	6.0	WY	7.5	7.9

Table 3: Nash subsidies

	Nash subsidy		Δ welfare		
	subsidy (%)	subsidy ($bn)	local (%)	local ($bn)	national (%)
AL	10.3	8.7	-1.2	-0.5	-1.3
AZ	11.3	7.6	-1.1	-0.3	-1.3
AR	9.0	4.4	-0.4	-0.1	-1.3
CA	12.4	65.7	-0.6	-1.3	-1.3
CO	10.8	5.2	-1.0	-0.2	-1.3
CT	9.9	6.8	-1.8	-0.6	-1.3
DE	7.4	1.1	-3.1	-0.2	-1.3
FL	11.3	12.9	-0.6	-0.3	-1.3
GA	9.4	12.1	-1.1	-0.7	-1.3
ID	8.5	1.0	1.1	0.1	-1.3
IL	8.5	18.8	-1.4	-1.5	-1.3
IN	9.1	17.7	-2.1	-2.1	-1.3
IA	10.5	9.3	-1.8	-0.8	-1.3
KS	9.5	5.4	-1.1	-0.3	-1.3
KY	8.0	6.4	-1.8	-0.7	-1.3
LA	11.7	8.7	-3.0	-1.0	-1.3
ME	10.1	1.8	-1.3	-0.1	-1.3
MD	6.7	2.7	-1.8	-0.4	-1.3
MA	10.4	10.8	-1.6	-0.8	-1.3
MI	10.6	22.1	-1.8	-1.7	-1.3
MN	10.8	11.3	-0.4	-0.2	-1.3
MS	8.4	3.3	-0.4	-0.1	-1.3
MO	9.4	9.7	-1.7	-0.8	-1.3
MT	5.4	0.1	2.3	0.0	-1.3
NE	8.3	2.2	-0.3	0.0	-1.3
NV	6.9	1.2	-0.3	0.0	-1.3
NH	6.5	1.2	0.4	0.0	-1.3
NJ	7.4	7.3	-1.6	-0.8	-1.3
NM	6.2	0.2	0.3	0.0	-1.3
NY	9.6	17.2	-1.8	-1.6	-1.3
NC	10.8	24.2	-2.4	-2.6	-1.3
ND	8.1	0.6	-0.9	0.0	-1.3
OH	9.5	24.9	-1.6	-2.0	-1.3
OK	10.2	4.2	-1.2	-0.2	-1.3
OR	11.5	10.0	-1.5	-0.6	-1.3
PA	9.1	19.4	-1.6	-1.7	-1.3
RI	6.1	0.6	0.7	0.0	-1.3
SC	8.2	6.1	-1.4	-0.5	-1.3
SD	8.6	0.9	-0.5	0.0	-1.3
TN	5.4	5.6	-0.8	-0.4	-1.3
TX	12.0	47.8	-1.9	-3.3	-1.3
UT	10.4	4.1	-0.5	-0.1	-1.3
VT	8.1	0.5	-0.6	0.0	-1.3
VA	9.8	10.0	-2.0	-1.0	-1.3
WA	11.7	13.2	-1.1	-0.6	-1.3
WV	6.1	1.0	-1.6	-0.2	-1.3
WI	10.4	17.1	-1.0	-0.7	-1.3
WY	7.0	0.2	-2.5	0.0	-1.3
Average	9.1	9.9	-1.1	-0.6	-1.3

Table 4: Sensitivity checks for Nash subsidies

Sensitivity wrt. sigma

	subsidy	Δ welfare	
σ	avg.	local	national
0.80	9.1	-1.1	-1.3
1.00	9.1	-1.1	-1.3
1.20	9.1	-1.1	-1.3
1.40	9.1	-1.1	-1.3
1.60	9.1	-1.1	-1.3

Sensitivity wrt. epsilon

	subsidy	Δ welfare	
ε	avg.	local	national
4.00	11.7	-2.8	-3.2
4.50	10.2	-1.7	-2.0
5.00	9.1	-1.1	-1.3
5.50	8.2	-0.8	-1.0
6.00	7.5	-0.6	-0.7

Sensitivity wrt. phi

	subsidy	Δ welfare	
ϕ	avg.	local	national
0.33	14.9	-4.5	-4.9
0.29	11.7	-2.2	-2.5
0.25	9.1	-1.1	-1.3
0.22	7.0	-0.6	-0.8
0.20	5.3	-0.3	-0.4

Sensitivity wrt. intial subsidies

state	min	max	state	min	max
AL	10.0	10.4	NE	8.0	8.4
AZ	11.1	11.4	NV	6.6	7.1
AR	8.6	9.0	NH	6.2	6.6
CA	12.4	12.5	NJ	7.1	7.5
CO	10.5	10.9	NM	6.2	6.5
CT	9.6	10.0	NY	9.4	9.8
DE	7.1	7.5	NC	10.6	10.9
FL	11.1	11.3	ND	7.8	8.2
GA	9.1	9.5	OH	9.3	9.6
ID	8.2	8.6	OK	10.0	10.4
IL	8.3	8.6	OR	11.2	11.6
IN	8.9	9.2	PA	8.9	9.2
IA	10.3	10.6	RI	5.8	6.2
KS	9.2	9.6	SC	8.0	8.4
KY	7.8	8.1	SD	8.3	8.7
LA	11.5	11.8	TN	5.1	5.4
ME	9.8	10.2	TX	11.9	12.0
MD	6.4	6.8	UT	10.1	10.5
MA	10.2	10.5	VT	8.0	8.4
MI	10.4	10.7	VA	9.5	9.8
MN	10.5	10.8	WA	11.5	11.8
MS	8.1	8.5	WV	5.9	6.2
MO	9.1	9.4	WI	10.2	10.5
MT	5.2	5.5	WY	6.7	7.1

Table 5: Cooperative subsidies

	Δ welfare w/ transfer			Δ welfare w/o transfers		
State	national (%)	local (%)	local ($bn)	national (%)	local (%)	local ($bn)
AL	0.5	4.8	1.9	0.00	-0.07	-0.03
AZ	0.5	-2.3	-0.7	0.00	0.06	0.02
AR	0.5	11.7	2.7	0.00	0.05	0.01
CA	0.5	3.5	7.8	0.00	0.03	0.06
CO	0.5	1.2	0.3	0.00	0.20	0.04
CT	0.5	-1.8	-0.6	0.00	-0.06	-0.02
DE	0.5	-16.9	-1.3	0.00	0.19	0.01
FL	0.5	7.1	3.6	0.00	0.01	0.01
GA	0.5	4.0	2.5	0.00	0.08	0.05
ID	0.5	19.5	1.1	0.00	0.13	0.01
IL	0.5	0.0	0.0	0.00	0.01	0.01
IN	0.5	-3.8	-3.7	0.00	-0.20	-0.19
IA	0.5	-4.6	-1.9	0.00	0.04	0.02
KS	0.5	3.4	0.9	0.00	0.01	0.00
KY	0.5	-1.8	-0.7	0.00	-0.02	-0.01
LA	0.5	-15.8	-5.4	0.00	-0.32	-0.11
ME	0.5	4.6	0.4	0.00	-0.12	-0.01
MD	0.5	-3.0	-0.7	0.00	0.13	0.03
MA	0.5	0.1	0.1	0.00	-0.02	-0.01
MI	0.5	-0.3	-0.3	0.00	-0.16	-0.16
MN	0.5	7.7	3.6	0.00	0.20	0.09
MS	0.5	11.7	2.3	0.00	-0.04	-0.01
MO	0.5	-1.5	-0.7	0.00	0.06	0.03
MT	0.5	29.8	0.4	0.00	0.10	0.00
NE	0.5	10.1	1.3	0.00	0.00	0.00
NV	0.5	-1.7	-0.2	0.00	0.14	0.01
NH	0.5	17.1	1.6	0.00	-0.04	0.00
NJ	0.5	-0.5	-0.3	0.00	-0.05	-0.03
NM	0.5	13.6	0.3	0.00	-0.56	-0.01
NY	0.5	0.6	0.5	0.00	-0.23	-0.20
NC	0.5	-8.9	-9.3	0.00	0.10	0.10
ND	0.5	0.5	0.0	0.00	0.22	0.01
OH	0.5	0.3	0.3	0.00	0.01	0.02
OK	0.5	7.9	1.5	0.00	-0.84	-0.16
OR	0.5	-9.8	-3.8	0.00	0.26	0.10
PA	0.5	1.4	1.5	0.00	0.05	0.05
RI	0.5	21.3	1.1	0.00	-0.10	-0.01
SC	0.5	2.4	0.9	0.00	-0.21	-0.08
SD	0.5	5.6	0.3	0.00	0.21	0.01
TN	0.5	2.3	1.3	0.00	0.07	0.04
TX	0.5	-7.7	-13.5	0.00	0.02	0.03
UT	0.5	5.2	0.9	0.00	0.14	0.03
VT	0.5	18.4	0.5	0.00	-0.88	-0.03
VA	0.5	-3.7	-1.8	0.00	0.18	0.09
WA	0.5	-4.7	-2.3	0.00	0.29	0.14
WV	0.5	-1.4	-0.1	0.00	0.00	0.00
WI	0.5	3.8	2.9	0.00	0.17	0.13
WY	0.5	-18.7	-0.2	0.00	0.19	0.00
Mean	0.5	2.3	-0.1	0.00	-0.01	0.00

Table 6: Sensitivity checks for cooperative subsidies

Sensitivity wrt. sigma

	subsidy	Δ welfare	
σ	avg.	local	national
0.80	0.0	2.7	0.5
1.00	0.6	2.9	0.5
1.20	0.0	2.3	0.5
1.40	0.0	2.2	0.5
1.60	0.0	2.0	0.5

Sensitivity wrt. epsilon

	subsidy	Δ welfare	
ε	avg.	local	national
4.00	0.0	3.6	0.8
4.50	0.0	2.8	0.6
5.00	0.0	2.3	0.5
5.50	0.0	2.0	0.5
6.00	0.0	1.8	0.4

Sensitivity wrt. phi

	subsidy	Δ welfare	
ϕ	avg.	local	national
0.33	0.0	2.9	0.8
0.29	0.0	2.5	0.6
0.25	0.0	2.3	0.5
0.22	0.0	2.1	0.5
0.20	0.9	2.4	0.4

Table 7: Local welfare weights

State	Weight (%)	State	Weight (%)
IN	0.54	MS	0.05
NY	0.52	GA	0.05
CA	0.41	KS	0.05
OK	0.40	RI	0.04
SC	0.38	AZ	0.04
MI	0.37	ME	0.03
IL	0.29	MD	0.03
TX	0.20	TN	0.03
NJ	0.20	OR	0.02
NM	0.19	WI	0.02
OH	0.17	UT	0.02
PA	0.16	ID	0.01
VT	0.15	MN	0.01
AL	0.14	VA	0.01
KY	0.12	WA	0.01
LA	0.11	NV	0.00
NC	0.10	AR	0.00
FL	0.10	MT	0.00
MA	0.09	NH	0.00
IA	0.08	ND	0.00
CT	0.08	CO	0.00
MO	0.06	SD	0.00
WV	0.05	DE	0.00
NE	0.05	WY	0.00