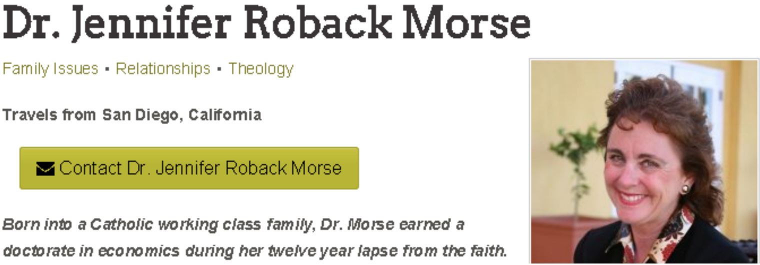
Introduction: Public Goods in General Equilibrium

- Why does San Francisco have higher wages and home prices than New Haven does?
- Hedonic models accounted for differences across neighborhoods
  - But wages don't vary across neighborhoods
- How do people choose what city to live in?
- If no congestion, everyone would live in location(s) with best amenities and productivity
- In reality, a fixed factor (e.g., land) imposes constraints.

Introduction: Public Goods in General Equilibrium

- Roback a famous family of models. Use her words:
  - This study focuses on the role of wages and rents in allocating workers to locations with various quantities of amenities.



– identical consumers

– equilibrium in rental & wage markets

$\Rightarrow V(w,r;s) = k$ everywhere given a w, r — consumers indifferent between locations

– CRS in firms $\Rightarrow$ can switch between cost & production function

$X = f(l^p, N; s)$

$C(w,r;s) = 1$ since equil. & competitive $\Rightarrow \pi = 0$ everywhere

$\Rightarrow$ firms are indifferent between locations

Introduction: Public Goods in General Equilibrium

- Roback a famous family of models. Use her words:
  - This study focuses on the role of wages and rents in allocating workers to locations with various quantities of amenities.
- General idea: locations vary in amenities and productivity
  - These attributes are exogenous
- How do land and labor markets clear?
- What does this reveal about demand for goods which vary across space?

Model: Setup

- s amenity $\in (S_1, S_2)$
- X Hicksian composite (price exogenous, numeraire)
- Capital and labor fully mobile
- No cost to commute within cities, infinite cost to commute between
- All workers identical

– recent Roback model papers have an extra term in the utility (all workers get the same utility from X except one term – which is the migration cost)

Model: Consumer Problem

$$\max U\left(x, l^c; s\right) \text{ s.t. } w + I = x + l^c r$$

- x Quantity of composite good consumed
- s Quantity of amenity in chosen location
- l^c Residential land consumed
- w Wages
- r Rental payments
- I Nonlabor income (independent of location)

Model: Consumer Problem

$$\max U(x, l^c; s) \text{ s.t. } w + I = x + l^c r$$

- x Quantity of composite good consumed
- s Quantity of amenity in chosen location
- l^c Residential land consumed
- w Wages
- r Rental payments
- I Nonlabor income (independent of location)

Model: Consumer Problem

Market equilibrium for workers:

$$V(w, r; s) = k$$

- V Indirect utility
- $\partial V / \partial s > 0$ Assume s is good amenity
- Because all workers identical, wages and rents equalize utility in all occupied locations
    - Otherwise some workers want to move and it's not an equilibrium
    - This indifference condition is a key result and intuition for these models!

Model: Consumer Problem

Show this equation again because it's important

$$V\left(w, r; s\right) = k$$

- In an equilibrium with homogenous consumers, a consumer gets the same utility from any location.
  - Otherwise a consumer would want to move, and it wouldn't be an equilibrium
- Most applications have in mind heterogeneous consumers but a condition like this generally applies to a marginal consumer.

Model: Firm Problem

- Firms produce X produce from CRS technology:

$$X = f\left(l^p, N; s\right)$$

- l^p land used in production
- N workers in the city
- Firm problem: minimize costs s.t. production function
- Since f is CRS, work with unit cost function
- Equilibrium: price equals marginal cost (which equals average cost by CRS).
    - Recall: one product X, price equals one, so

$$C\left(w, r; s\right) = 1$$

- This is the firm's indifference condition.
- In equilibrium with homogeneous firms, firms get same profit from all locations

Model: Firm Problem

- State this again because it's important
- In equilibrium with homgeneous firms, firms get same profit from all locations

Model: Firm Problem

- Unproductive amenity: $C_s < 0$ → in equil., if we have more of an unproductive amenity, cost must be lower (but production less productive) (w, r must be lower b/c otherwise firms would move)
    - Strange idea: unproductive amenity
    - Example: clean air
- Example of productive amenity: lack of snow storms
- Example of amenity that doesn't affect productivity: sunny days

Model: Aside

- Worth noting respects in which this basic model is simple by being unrealistic
  - Homogenous firms and consumers
    - So basic model won't try to recover heterogeneous tastes and cost functions
    - No Rosen's second step here
  - No mobility costs
  - One product X
- This lets the model get at basic conclusions.
  - Later in the paper it adds some heterogeneity. Other papers do this also.
- Chekhov's gun:

  "If in the first act you have hung a pistol on the wall, then in the following one, it should be fired. Otherwise don't put it there

Model: Equilibrium

- The consumer and firm indifference conditions pin down w and r as functions of s, given k
- Assume unproductive s.
  - If $s_2 > s_1$, city 2 needs lower factor prices to equalize firm costs

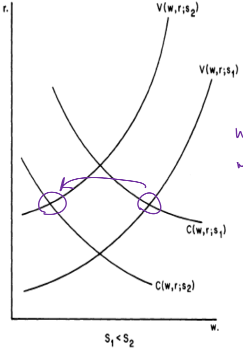
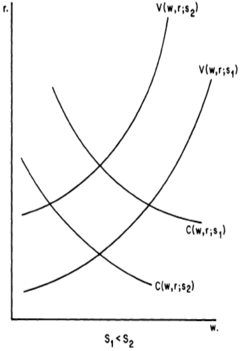


Fig. 1

when moving to a higher unproductive amenity, wages fall, but effects on <u>rent</u> is ambiguous

Model: Equilibrium

- In places with a greater level of unproductive amenity,
  - Lower wages but ambiguous change in rents
- Intuition: for unproductive s, firms prefer low s and workers prefer high s
  - Low wages along with high s equilibrate these conflicting demands
  - But both groups dislike rents so rental effect is ambiguous



Model: Equilibrium

- Algebraic version:
- Differentiate the firm and worker indifference conditions, solve for dw/ds and dr/ds

$$
\begin{aligned}
\frac{dw}{ds} &= \frac{1}{\Delta}\left(-V_s C_r + C_s V_r\right) < 0 \\
\frac{dr}{ds} &= \frac{1}{\Delta}\left(-V_w C_s + V_s C_w\right) \gtrless 0
\end{aligned}
$$

where $\Delta = V_w C_r - V_r C_w > 0$

$L(s)$ is total land at location s

- Implies $dw/ds < 0$ but dr/ds ambiguous sign.

Model: Equilibrium

- Combining and using Roy's identity gives

$$
\begin{aligned}
p_s^* &= \frac{V_s}{V_w} = l^c \frac{dr}{ds} - \frac{dw}{ds} \\
\frac{p_s^*}{w} &= k_l \frac{d \log r}{ds} - \frac{d \log w}{ds}
\end{aligned}
$$

- k_l is share of land in consumer budget
- Interpretation: value of small change in amenity s is increase in rents, weighted by land budget share, minus increase in wages
- Restated: value of amenity is amount of numeraire and residential land they sacrifice for it
- Applications:
  - Ranking cities
  - Benefit-cost analysis of city-wide crime, pollution, schools, etc.
  - Adjusting GNP by quality of life measures (Nordhaus and Tobin 1972)

Extension: Nontraded Goods

- y nontraded goods (housing, haircuts)
- Now consumption land l^c is input to housing production
- l^y land input to nontraded goods production
- p relative price of nontraded relative to traded goods

$$V(w, p; s) = k$$

- r doesn't matter directly for indirect utility now

Extension: Nontraded Goods

- Nontraded goods production:

$$G\left(w, r; s\right) = p\left(s\right)$$

- Market clearing: total output of nontraded goods equals Ny
- Again, consumer and firm indifference conditions determine w, r, and p.

Extension: Nontraded Goods

- Similar approach as before gives

$$p_s^* = y\frac{dp}{ds} - \frac{dw}{ds}$$

- Again p_s^* is change in numeriare income required for small variation in s.
- One implication: easier to have comparative statics for land prices than for home prices

Model: Comments

- How can such a simple model have nontrivial insights?
- Sign of a good model, more common in 70s/80s and in price theory

Empirical Application

- Now assume consumers are heterogeneous.
  - [If empirical application requires this, why didn't the theory incorporate it?]
- People with stronger preference for amenity sort into more amenable places, accept lower wage
- So marketwide estimates overstate utility costs for some people, understate for others

Empirical Application

- Type A: strong preference for amenity
- Type B: weak preference for amenity.
- We observe A and B.
  - But we should compare A and A', or compare B and B'

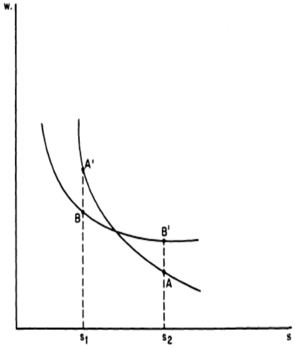


FIG. 2

Empirical Application

- This is like the last section of an econometrica article–demonstrating an application of the theory, but not a stand-alone paper
- Goal: rank quality of life across cities
  - Remember: utility needs to be equal across cities
  - So what is quality of life? It's the aggregate of amenities—everything excluding wages and home prices

seek this information from the data. Viewed another way, theory does not tell us which attributes are goods; theory only tells us how people behave with respect to goods. For this reason, extensive experimentation was done with many variables. Second, the small number of observations and high degree of multicollinearity among the variables limited the number of different indicators which could be used. For instance, when several climate variables are entered into the same earnings regression, none is significant. Yet one might have thought that both the number of hot days and the precipitation would be relevant to location decisions. Also, some of the results are sensitive to alternative specifications. The equations reported below were chosen to be representative of the bulk of the results and to be demonstrative of the type of behavior described by the theory.

Empirical Application

- Data: May 1973 CPS (98 Cities)
  - Men over 18 with positive earnings
- FHA Homes: site prices per square foot for 83 of 98 cities
  - Overrepresents low-income families
  - No information on home location within a city

Because of these limitations of the data, the land price results presented below are merely intended to be illustrative of the method outlined in the theory

Empirical Application: Individual Earnings Results

Regression of Log of Weekly Earnings on Personal Characteristics

	Coefficient	t-Statistic
Intercept	3.7418	105.60
Household head	.1457	9.94
White	.0272	2.14
Married	.0776	6.86
Veteran	.0274	3.07
School	.0446	28.52
Experience	.0285	26.57
Experience squared	−.0005	−24.85
Hours	.0101	20.42
Part time	−.2869	−17.90
Private	.0129	.89
Professional	.3263	24.48
White collar	.1189	8.51
Blue collar	.1092	9.52
Poverty incidence	−.9063	−18.05
Construction	.1333	6.19
Durables	−.0519	−2.52
Nondurables	−.0589	−2.69
Transport	.0192	.90
Trade	−.1463	−7.22
Services	−.2085	−11.49
Union	.1213	14.70

NOTE.—Data are from the May 1973 Current Population Survey; R^2 = .4881; F-ratio = 543.9; N = 12,001. The omitted occupation is laborers; the omitted industry is public administration.

Empirical Application: City Wages and Amenities

TABLE 1

COEFFICIENTS OF CITY CHARACTERISTICS FROM LOG EARNINGS REGRESSIONS IN 98 CITIES

	1	2	3	4
TCRIME 73	$.94 \times 10^{-5}$	$.44 \times 10^{-5}$	$.74 \times 10^{-5}$	$.86 \times 10^{-5}$
	(2.58)	(1.17)	(1.93)	(2.21)
UR 73	$.36 \times 10^{-2}$	$.12 \times 10^{-2}$	$.32 \times 10^{-2}$	$.27 \times 10^{-2}$
	(1.29)	(.43)	(1.14)	(.97)
PART 73	$.24 \times 10^{-3}$	$.13 \times 10^{-3}$	$.37 \times 10^{-3}$	$.34 \times 10^{-3}$
	(1.55)	(.86)	(2.33)	(2.15)
POP 73	$.16 \times 10^{-7}$	$.15 \times 10^{-7}$	$.16 \times 10^{-7}$	$.16 \times 10^{-7}$
	(7.97)	(7.74)	(8.04)	(8.11)
DENSSMSA	$.81 \times 10^{-6}$	$.24 \times 10^{-5}$	$.20 \times 10^{-5}$	$.38 \times 10^{-5}$
	(.29)	(.86)	(.73)	(1.40)
GROW 6070	$.21 \times 10^{-2}$	$.14 \times 10^{-2}$	$.15 \times 10^{-2}$	$.17 \times 10^{-2}$
	(7.84)	(5.66)	(6.06)	(6.47)
HDD	$.20 \times 10^{-4}$			
	(8.48)			
TOTSNOW		$.72 \times 10^{-3}$		
		(3.54)		
CLEAR			$-.64 \times 10^{-2}$	
			(−4.80)	
CLOUDY				$.72 \times 10^{-2}$
				(5.21)
R^2	.4980	.4955	.4960	.4962
F-ratio	424.2	420.0	420.8	421.1
N = 12,001				

NOTE.—Regressions include all personal characteristics. Sample includes 98 cities; t-statistics are in parentheses (see App. for variable definitions).

Empirical Application: City Wages and Amenities

TABLE 2

COEFFICIENTS OF REGION DUMMIES AND CITY CHARACTERISTICS

NORTHEAST	−.0218	−.0095
	(−2.25)	(−.74)
SOUTH	−.0669	−.0138
	(−6.51)	(−.87)
WEST	−.0354	−.0579
	(−3.46)	(−3.41)
TCRIME 73		$.13 \times 10^{-4}$
		(2.82)
UR 73		$.92 \times 10^{-2}$
		(2.60)
PART 73		$.29 \times 10^{-3}$
		(1.87)
POP 73		$.16 \times 10^{-7}$
		(7.77)
DENSSMSA		$-.13 \times 10^{-5}$
		(−.42)
GROW 6070		$.23 \times 10^{-2}$
		(8.41)
HDD		$.16 \times 10^{-4}$
		(4.86)
R^2	.4900	.4986
F-ratio	479.4	384.0

NOTE.—Regressions include all personal characteristics. Sample includes all 98 cities; t-statistics are in parentheses.

Empirical Application: City Land Values and Amenities

TABLE 3

REGRESSIONS OF THE LOG OF AVERAGE RESIDENTIAL SITE PRICE PER SQUARE FOOT ON CITY CHARACTERISTICS

	1	2	3	4
TCRIME 73	2.5×10^{-5}	1.5×10^{-5}	-4.5×10^{-7}	7.0×10^{-6}
	(.65)	(.38)	(−.01)	(.16)
UR 73	8.9×10^{-2}	8.8×10^{-2}	9.2×10^{-2}	9.1×10^{-2}
	(3.45)	(3.35)	(3.53)	(3.52)
PART 73	2.2×10^{-4}	1.1×10^{-4}	-3.8×10^{-5}	1.4×10^{-4}
	(.15)	(.08)	(−.02)	(.09)
POP 73	6.8×10^{-8}	6.9×10^{-8}	6.8×10^{-8}	6.8×10^{-8}
	(1.80)	(1.78)	(1.76)	(1.76)
DENSSMSA	1.9×10^{-4}	2.0×10^{-4}	2.0×10^{-4}	2.0×10^{-4}
	(3.02)	(3.12)	(3.17)	(3.18)
GROW 6070	1.1×10^{-2}	1.0×10^{-2}	9.9×10^{-3}	1.0×10^{-2}
	(4.34)	(4.11)	(4.03)	(4.00)
HDD	3.5×10^{-5}			
	(1.44)			
TOTSNOW		1.3×10^{-3}		
		(.69)		
CLEAR			1.2×10^{-4}	
			(.09)	
CLOUDY				3.2×10^{-4}
				(.21)
INTERCEPT	−1.73	−1.54	−1.44	−1.53
	(−5.92)	(−5.99)	(−6.51)	(−3.32)
R^2	.5741	.5650	.5623	.5625
F-ratio	14.44	13.92	13.77	13.78

SOURCE.—Data are from U.S. Department of Housing and Urban Development 1973. $N = 83$.

Empirical Application: Implicit Amenity Prices

TABLE 4

IMPLICIT PRICES OF AMENITIES COMPUTED FROM TABLES 1 AND 3

	1	2	3	4
TCRIME 73				
(crimes/100 population)	$−9.25	$ .90	$ −8.05	$ −9.15
UR 73				
(fraction unemployed)	−5.55	20.65	−.70	5.00
PART 73				
(micrograms/cubic meter)	−2.50	−1.40	−4.00	−3.70
POP 73				
(10,000 persons)	−1.50	−1.40	−1.50	−1.50
DENSSMSA				
(100 persons/square mile)	6.30	4.90	5.35	3.35
GROW 6070				
(percentage change in population)	−1.85	−11.95	−13.05	−15.2
HDD				
(1° F colder for one day	−.20			
TOTSNOW				
(inches)		−7.30		
CLEAR				
(days)			69.55	
CLOUDY				
(days)				−78.25

NOTE.—Measurement units of amenities shown under variable name. Each entry is computed using eq. (5) in the text and evaluated at mean annual earnings. $p_s^* = [k_l(d \log r/ds) - (d \log w/ds)]w$. Average annual earnings = $10,868. Average budget share of land = .035. Negative numbers indicate disamenities, while positive numbers indicate amenities.

Empirical Application: Quality of Life Measures

TABLE 5

RANK CORRELATIONS BETWEEN VARIOUS MEASURES OF QUALITY OF LIFE INDEX

	QOL 1	QOL 2	QOL 3	QOL 4
QOL 1	1.000	.7846	.2480	.2902
	(.0)	(.0001)	(.0138)	(.0037)
QOL 2		1.000	.3568	.2701
		(.0)	(.0003)	(.0072)
QOL 3			1.000	.8219
			(.0)	(.0001)
QOL 4				1.000
				(.0)

NOTE.—Probabilities in parentheses. The indices are computed from cols. 1–4 of tables 2 and 6. Thus QOL 1 uses HDD, QOL 2 uses TOTSNOW, QOL 3 uses CLEAR, and QOL 4 uses CLOUDY.

Empirical Application: Quality of Life Measures

TABLE 6

COMPARISON OF QOL 3 RANKINGS OF 20 LARGEST CITIES WITH RANKING OF LIU

Rank	Name	Liu's Rank	QOL 3	Population Rank
1	Los Angeles–Long Beach	10	1.7517	2
2	Anaheim–Santa Ana–Garden Grove	9	1.7363	19
3	San Francisco–Oakland	2	1.5841	6
4	Dallas	5	1.3378	17
5	Baltimore	13	1.0244	12
6	Nassau-Suffolk	. . .	1.0010	9
7	St. Louis	16	.9407	11
8	Milwaukee	8	.9386	20
9	Boston	12	.9296	8
10	Minneapolis	4	.9047	16
11	New York	14	.8962	1
12	Washington, D.C.	3	.8910	7
13	Newark	11	.8853	15
14	Philadelphia	7	.8038	4
15	Houston	6	.7708	14
16	Chicago	18	.7416	3
17	Detroit	17	.6347	5
18	Cleveland	15	.6227	13
19	Seattle-Everett	1	.5871	18
20	Pittsburgh	19	.4961	10

NOTE.—Liu's rank is based on adjusted standardized score of environmental component. Nassau-Suffolk is not included in Liu's (1976) study.

Roback: Conclusions

- Accounting for land and labor markets simultaneously straightforward
- Model looks fairly simple, empirical applications are sensible
- But, this inspired a wide range of applications (we'll now discuss a few)
- One important assumption: ability to measure land prices
  - If using home prices, assume inelastic housing supply or otherwise account for supply elasticity.

Introduction: Short Run v. Long Run

- Many empirical papers estimate short-run relationships
  - Reduced-form regressions: identifying variation often more credible in short run
  - Differences-in-differences, RD, etc.
  - Estimate partial equilibrium effects (esp. with fixed effects)
- Economic models more often about long-run relationships
  - Comparative statics
  - General equilibrium
  - Growth, dynamic models
- How reconcile the two?

Introduction: Local v. National Relationships

- Many empirical papers estimate local relationships
  - Other regions provide comparison groups
  - DD, RD
- This makes it difficult to estimate spillovers or national effects of a local policy

Introduction: This Paper

- Estimate national, long-term consequences of a place-based policy
- Tennessee Valley Authority (TVA)
  - Part of New Deal programs
  - Federal subsidies between 1933-1960
  - In 1950-1955, subsidy for mean household was $750/year (10% of household income)
- Two methodologies
  - Reduced-form policy evaluation: compare TVA against other proposed authorities
  - Dynamic panel model

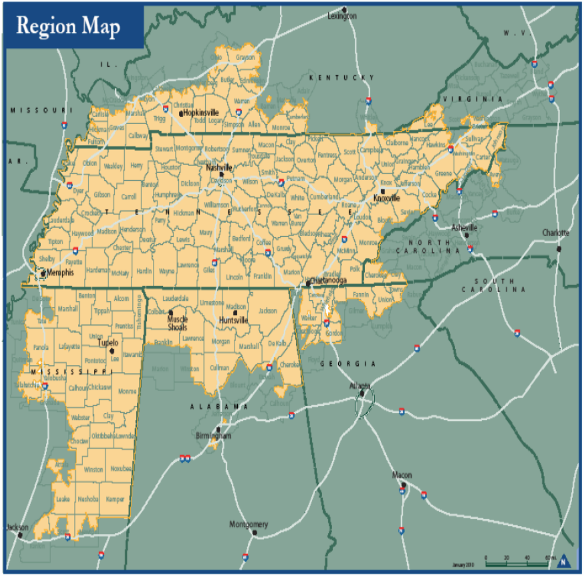
Introduction: This Paper

- General comments
  - If this paper was written as "DD policy evaluation of TVA," it would have been less interesting
  - Writing it as "DD policy evaluation of long-run local effects of place-based policy" is more interesting.
  - Writing it as "DD policy evaluation of long-run NATIONAL effects of place-based policy" makes it more interesting.
  - Comparing reduced-form DD and distinct dynamic panel model estimates makes it an even better paper.

Background: TVA

- Substance: hydroelectric dams; navagation canal; roads; new schools; flood control
  - Other minor programs: malaria prevention, soil erosion mitigation, education, health clinics, cheap fertilizer, reforestation, forest fire control, advising for local development
- Electricity: designed to bring in manufacturing industries to an agricultural region.
- Total appropriations $20 billion 1934-2000
  - 73% of this was 1940-1958
- 1959: federal legislation makes it self-financing

Background: TVA



Research Design: DD

- Selection criteria for counties into program
  - Rural, needed electricity
  - Flooding and/or misguided land use
  - Deficits
  - Lacked public facilities (libraries, health services, schools)
  - Willing to receive advice from TVA
  - Had planning agencies, enabling legislation, willing to try new fertilizers
  - Within transmission distance of power plants

Research Design: DD

Table I: Summary Statistics

	Overall				Trimmed Sample	
	TVA	Non-TVA	Non-TVA South	Non-TVA Proposed Authorities	Non-TVA	Non-TVA South
	(1)	(2)	(3)	(4)	(5)	(6)
1930 Characteristics						
Log Population	9.991	9.977	9.989	9.940	9.905	9.979
Log Employment	8.942	8.967	8.959	8.908	8.881	8.947
Log # of Houses	8.445	8.508	8.455	8.466	8.442	8.445
Log Average Manufacturing Wage	1.406	1.802	1.545	1.685	1.728	1.538
Manufacturing Employment Share	0.075	0.090	0.080	0.077	0.080	0.078
Agricultural Employment Share	0.617	0.455	0.541	0.510	0.487	0.547
% White	0.813	0.885	0.722	0.830	0.863	0.724
% Urbanized	0.153	0.280	0.233	0.216	0.242	0.215
% Illiterate	0.088	0.045	0.092	0.060	0.051	0.092
% of Whites Foreign Born	0.002	0.059	0.013	0.020	0.030	0.011
Log Average Farm Value	5.252	5.646	5.386	5.552	5.579	5.370
Log Median Housing Value	9.271	9.581	9.360	9.452	9.516	9.358
Log Median Contract Rent	8.574	9.030	8.679	8.834	8.934	8.672
% Own Radio	0.079	0.296	0.114	0.210	0.256	0.112
Max Elevation (meters)	1576.190	2364.531	1068.943	1758.893	2044.656	1070.334
Elevation Range (Max-Min)	1127.761	1521.322	712.336	1083.293	1251.074	715.253
% Counties in South	1.000	0.342	1.000	0.554	0.447	1.000

Research Design: DD

Changes 1920-1930						
Log Population	0.051	0.049	0.067	0.004	0.037	0.060
Log Employment	0.082	0.096	0.111	0.045	0.083	0.103
Log # of Houses	0.078	0.092	0.108	0.046	0.078	0.100
Log Average Manufacturing Wage	0.117	0.217	0.108	0.172	0.197	0.103
Manufacturing Employment Share	-0.010	-0.035	-0.018	-0.018	-0.026	-0.018
Agricultural Employment Share	-0.047	-0.036	-0.047	-0.046	-0.042	-0.047
% White	0.012	-0.011	-0.010	0.000	-0.006	-0.004
% Urbanized	0.047	0.064	0.080	0.042	0.054	0.069
% Illiterate	-0.030	-0.014	-0.029	-0.019	-0.015	-0.028
% of Whites Foreign Born	-0.001	-0.023	-0.016	-0.012	-0.015	-0.012
Log Average Farm Value	-0.013	-0.076	0.025	-0.182	-0.102	0.013
# of Observations	163	2326	795	828	1744	779
# of States	6	46	14	25	43	14

Notes: The unit of observation is a county. The trimmed sample is obtained by dropping control counties which, based on their pre-program characteristics, have a predicted probability of treatment in the bottom 25 percent. All monetary values are in constant 2000 dollars. Data are from the 1920 and 1930 Census of Population and Housing, with the exception of farm value data, which are from the 1920 and 1930 Agricultural Census, and elevation data, which were collected by Fishback, Haines, and Kantor (2011). Manufacturing wage is obtained by dividing the total annual wage bill in manufacturing by the estimated number of workers in the industry. Details on data construction and limitations are provided in the online Appendix.

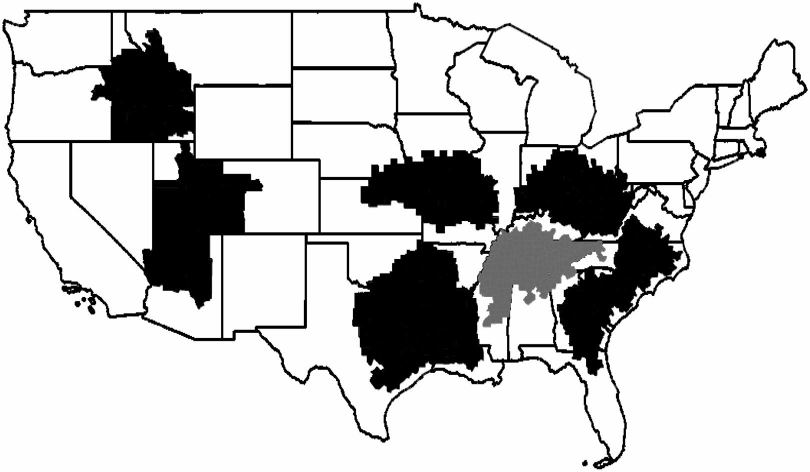
Research Design: DD

- Federal legislators proposed six other authorities modeled on TVA
  - Never actually implemented
  - Legislation didn't specify exact locations of six other authorities
- They create an algorithm to define scope of other authorities
  - Define all possible sets of spatially adjacent counties within a census region
  - For western authority, exclude counties in an existing authority (Bonneville Power Administration)
  - Keep all subregions with area equal to a third of total area in the region
  - Take 1920 and 1930 variables to proxy TVA eligibility: urban share; agricultural employment share; illiteracy; manufacturing wage and share; population; employment; mean farm value; median housing value; median rent
  - Select authority that minimizes Euclidean distance between subregion's vector and TVA's vector

Research Design: DD

- Note: this is an appealing research design in general
  - Instead of comparing all treated to all control units,
  - compare treated units to control units that were "in the pipeline" (applied but not yet received) or applied and rejected
  - This addresses selection on motivation
  - But selection concerns may remain

Research Design: DD



Methodology: DD

Oaxaca-Blinder regressions. First estimate for non-TVA counties

$$y_{it} - y_{it-1} = \alpha + \beta X_i + (\varepsilon_{it} - \varepsilon_{it-1})$$

- y_{it} dependent variable in county i, year t
- X_i preprogram characteristics
- Then use $\widehat{\beta}$ to predict counterfactual mean for TVA counties
- X_i include 38 variables measured in 1920 and 1930.
  - Quadratic in 1920, 1930 log population and interactions; 1920 and 1930 urban share, log employment; etc.
- Remove counties that border TVA region

Methodology: DD

Oaxaca-Blinder regressions. First estimate for non-TVA counties

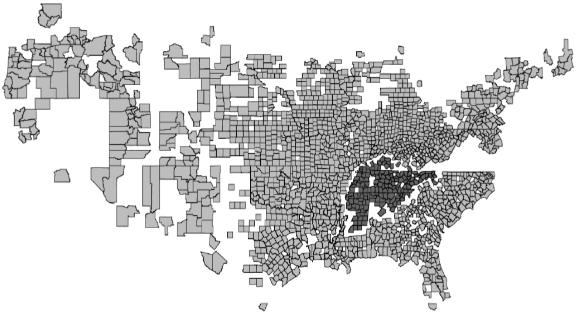
$$y_{it} - y_{it-1} = \alpha + \beta X_i + (\varepsilon_{it} - \varepsilon_{it-1})$$

- Three comparison groups
  - Rest of U.S.
  - Rest of South
  - Proposed Authorities
- Beenfit of Oaaca-Blinder regressions
  - Estimates ATT in presence of treatment effect heterogeneiety
- Can also be interpreted as propensity score reweighting estimator

Methodology: DD

- Also, drop counties which look substantially different from TVA counties
- Estimate logit of probability of being in TVA
  - Drop counties with predicted treatment probability in bottom 25 percent
- Spatial autocorrelation
  - Cluster standard errors by state
  - Also use Conley (1999) standard errors ("HAC")

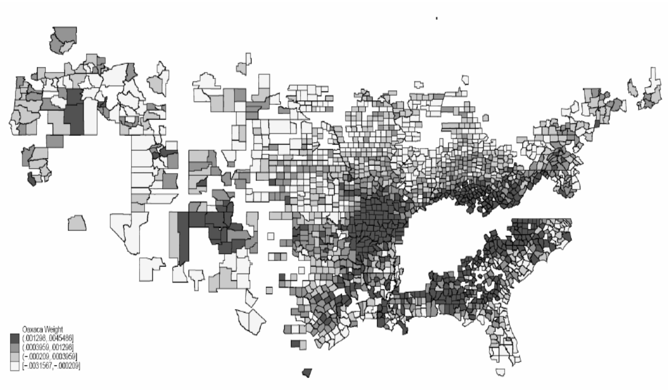
Methodology: DD



Shaded Areas are in TVA

Methodology: DD

Implicit weights on comparison counties



Results: 1900-1940 Placebo Test

Table II: Decadalized Growth Rates in TVA Region vs. Conterfactual Regions 1900-1940

	Outcome	Point Estimate (Unadjusted)	Clustered S.E.	Point Estimate (Controls)	Clustered S.E.	Spatial HAC	N
		(1)	(2)	(3)	(4)	(5)	(6)
	PANEL A: TVA Region vs Rest of US						
(1)	Population	0.007	(0.016)	0.010	(0.012)	(0.016)	1776
(2)	Total Employment	-0.009	(0.016)	0.005	(0.013)	(0.016)	1776
(3)	Housing Units	-0.006	(0.015)	0.007	(0.011)	(0.013)	1776
(4)	Average Manufacturing Wage	0.009	(0.018)	0.010	(0.021)	(0.016)	1428
(5)	Manufacturing Share	0.007*	(0.004)	0.005	(0.004)	(0.005)	1776
(6)	Agricultural Share	-0.007*	(0.004)	-0.001	(0.005)	(0.005)	1776
(7)	Average Agricultural Land Value	0.078***	(0.021)	0.025	(0.018)	(0.018)	1746

Results: 1900-1940 Placebo Test

	PANEL B: TVA Region vs. U.S. South					
(1)	Population	-0.018	(0.018)	0.003	(0.016)	850
(2)	Total Employment	-0.028	(0.018)	0.001	(0.016)	850
(3)	Housing Units	-0.025	(0.016)	0.005	(0.013)	850
(4)	Average Manufacturing Wage	0.001	(0.015)	0.001	(0.016)	687
(5)	Manufacturing Share	0.005	(0.005)	0.005	(0.005)	850
(6)	Agricultural Share	0.003	(0.004)	-0.002	(0.005)	850
(7)	Average Agricultural Land Value	-0.009	(0.020)	-0.007	(0.017)	839

Results: 1900-1940 Placebo Test

	PANEL C: TVA Region vs. Proposed Authrorities					
(1)	Population	0.026	(0.019)	0.011	(0.016)	926
(2)	Total Employment	-0.012	(0.017)	0.006	(0.015)	926
(3)	Housing Units	-0.014	(0.016)	0.006	(0.013)	926
(4)	Average Manufacturing Wage	0.012	(0.015)	0.008	(0.017)	734
(5)	Manufacturing Share	0.007	(0.006)	0.005	(0.006)	926
(6)	Agricultural Share	-0.005	(0.006)	0.004	(0.006)	926
(7)	Average Agricultural Land Value	0.080***	(0.026)	0.017	(0.018)	908

Results: Long-Run DD

Table III: Decadalized Impact of TVA on Growth Rate of Outcomes (1940-2000)

	Outcome	Point Estimate (Unadjusted)	Clustered S.E.	Point Estimate (Controls)	Clustered S.E.	Spatial HAC	N
		(1)	(2)	(3)	(4)	(5)	(6)
	PANEL A: TVA Region vs Rest of US						
(1)	Population	0.004	(0.021)	0.007	(0.020)	(0.018)	1907
(2)	Average Manufacturing Wage	0.027***	(0.006)	0.005	(0.004)	(0.005)	1172
(3)	Agricultural Employment	-0.130***	(0.026)	-0.056**	(0.024)	(0.027)	1907
(4)	Manufacturing Employment	0.076***	(0.013)	0.059***	(0.015)	(0.023)	1907
(5)	Value of Farm Production	-0.028	(0.028)	0.002	(0.032)	(0.026)	1903
(6)	Median Family Income (1950-2000 only)	0.072***	(0.014)	0.021	(0.013)	(0.011)	1905
(7)	Average Agricultural Land Value	0.066***	(0.013)	-0.002	(0.012)	(0.016)	1906
(8)	Median Housing Value	0.040**	(0.017)	0.005	(0.015)	(0.015)	1906

Results: Long-Run DD

	PANEL B: TVA Region vs. U.S. South					
(1)	Population	-0.007	(0.018)	0.014	(0.019)	942
(2)	Average Manufacturing Wage	0.003	(0.006)	0.001	(0.005)	610
(3)	Agricultural Employment	-0.097***	(0.030)	-0.051*	(0.027)	942
(4)	Manufacturing Employment	0.079***	(0.023)	0.063***	(0.024)	942
(5)	Value of Farm Production	-0.005	(0.025)	-0.006	(0.026)	939
(6)	Median Family Income (1950-2000 only)	0.041***	(0.012)	0.024**	(0.011)	942
(7)	Average Agricultural Land Value	0.031*	(0.018)	-0.003	(0.017)	942
(8)	Median Housing Value	0.019	(0.017)	0.007	(0.016)	942

Results: Long-Run DD

	PANEL C: TVA Region vs. Proposed Authorities				
(1)	Population	0.011	(0.018)	0.001	(0.017)	991
(2)	Average Manufacturing Wage	0.018***	(0.007)	0.005	(0.006)	618
(3)	Agricultural Employment	-0.101***	(0.029)	-0.071***	(0.027)	991
(4)	Manufacturing Employment	0.066***	(0.024)	0.053**	(0.024)	991
(5)	Value of Farm Production	0.002	(0.026)	0.011	(0.035)	989
(6)	Median Family Income (1950-2000 only)	0.060***	(0.012)	0.025**	(0.011)	991
(7)	Average Agricultural Land Value	0.060***	(0.019)	-0.003	(0.016)	991
(8)	Median Housing Value	0.033**	(0.016)	0.009	(0.016)	991

Results: DD Summary

- Short-run increase in all employment
- Long-run increase in manufacturing, decrease in agriculture
  - Manufacturing continued expanding after subsidies stopped!
- Long-run increase in household income
- No change in wages or housing values

Results: DD Summary

- How reconcile with Roback model?
  - Roback assumes inelastic supply of land
  - But here, elastic supply of houses and workers
  - So wages and home prices may underestimate welfare effects
    - Though they find no change in population
  - Wage data are low quality (no controls for worker characteristics)
- This is crying out for an analysis of underlying economics
  - That's purpose of dynamic panel estimates

Dynamic Panel: Model

- Counties are small open economies
  - Price takers on capital, labor, output markets
- Homogenous workers, perfectly mobile
- Counties heterogeneous in three respects
  - Exogenous amenity differences
  - Exogenous productivity
  - Endogenous agglomeration

Dynamic Panel: Model

Mobile and homogenous workers implies worker indifference across locations

$$\ln w_{it} + M_{it} = \overline{u}_t$$

w_{it} wages
M_{it} amenity levels
$\overline{u}_t$ reservation utility

Dynamic Panel: Model

Firms have Cobb-Douglas technology

$$Y_{it} = A_{it} K_{it}^{\alpha} F_i^{\beta} L_{it}^{1-\alpha-\beta}$$

A_{it} local productivity
L_{it} manufacturing workers
K_{it} local capital stock
F_i fixed factor (e.g., land) so derived labor demand slopes downward each period

Dynamic Panel: Model

Manufacturing good is numeriare ($p = 1$).
Firm's FOC imply inverse labor demand curve:

$$\ln w_{it} = C - \frac{\beta}{1-\alpha} \ln L_{it} + \frac{\beta}{1-\alpha} \ln F_i - \frac{\alpha}{1-\alpha} \ln r_t + \frac{1}{1-\alpha} \ln A_{it}$$

r_t national price of capital
$C \equiv \ln(1 - \alpha - \beta) + \frac{\alpha}{1-\alpha} \ln \alpha$

Dynamic Panel: Model

Assumed productivity decomposition:

$$\ln A_{it} = g\left(\frac{L_{it-1}}{R_i}\right) + \delta_t D_i + \eta_i + \gamma_t + \varepsilon_{it}$$

$\ln A_i$ log productivity
L_{it} manufacturing workers
R_i land area
$g(\cdot)$ agglomeration function
D_i TVA treatment dummy
η_i county fixed effect
γ_t National productivity shocks
ε_{it} idiosyncratic component of county productivity

Dynamic Panel: Model

Productivity, rewritten in steady-state:

$$\ln A_i = g\left(\frac{L_i}{R_i}\right) + \eta_i + \delta D_i$$

Output, rewritten in steady-state:

$$\begin{aligned} \ln Y_i &= \frac{\alpha}{1-\alpha}\ln\alpha + \frac{1-\alpha-\beta}{1-\alpha}\ln L_i + \frac{\beta}{1-\alpha}\ln F_i - \frac{\alpha}{1-\alpha}\ln r \\ &\quad + \frac{1}{1-\alpha}\ln A_i \end{aligned}$$

These are pretty simple—just omitting time-varying terms

Dynamic Panel: Model

Impact of increasing productivity of TVAs investment on county output:

$$\frac{dY_i}{d\delta} = \frac{1}{1-\alpha} Y_i \left(D_i + \frac{1-\alpha-\beta+\sigma_i}{L_i} \frac{dL_i}{d\delta} \right)$$

where local agglomeratoin elasticity is

$$\sigma_i \equiv \frac{d \ln A_i}{d \ln \left(\frac{L_i}{R_i} \right)} = g' \left(\frac{L_i}{R_i} \right) \frac{L_i}{R_i}$$

So two effects of TVA:

- Raise output in treated areas by $\frac{1}{1-\alpha}$ percent
- Indirect effect through endogenous labor adjustment (adding manufacturing workers mechanically raises output and increases agglomeration effects)

Dynamic Panel: Model

So what is aggregate effect of TVA on national output?

- Direct effect: positive
- Indirect effect: depends on shape of agglomeration forces.
- Moving worker from county i to j increases total output iff

$$\frac{Y_i}{L_i}\left(1-\alpha-\beta+\sigma_i\right) < \frac{Y_j}{L_j}\left(1-\alpha-\beta+\sigma_j\right)$$

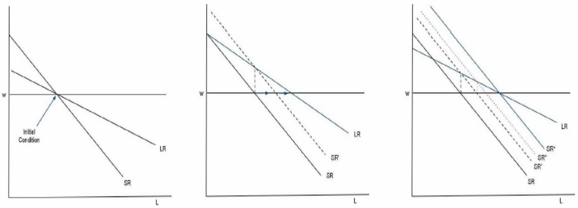
Dynamic Panel: Model

- If amenities are equal across two communities, wages are also equal, so mean labor productivity also equal
  - Then reallocation raises output only if agglomeration elasticity is greater in community j
  - If agglomeration elasticity same everywhere $(\sigma_i = \sigma)$, labor reallocation has no aggregate effects
- If amenities differ by community, then workers moving to low-amenity (high-wage) areas raises aggregate output
  - But workers have fewer amenities
  - Aggregate utility doesn't change

Dynamic Panel: Model

Compare aggregate output when agglomeration elasticity is constant v. non-constant

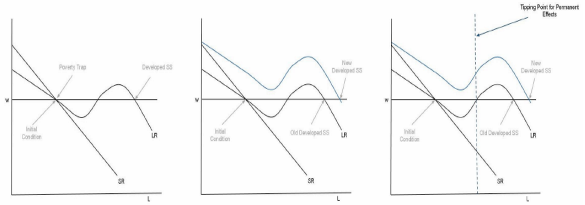
Figure IVa: Dynamics under Linear Agglomeration



x-axis is log manufacturing density, y-axis is log manufacturing wage, lines are inverse labor demand
LR: agglomeration dampens labor demand
First panel: initial steady-state
Second panel: outcomes with TVA

Dynamic Panel: Model

Figure IVb: Dynamics under Nonlinear Agglomeration



Now TVA can increase aggregate productivity, and has long-lasting effects
Key is multiple equilibria

Dynamic Panel: Model

Estimate agglomeration function with three piece linear spline

$$g\left(\frac{L_{it}}{R_i};\theta\right) = \sum_{k=1}^{3} \theta_k g_k\left(\frac{L_{it}}{R_i}\right)$$

where $g_k\left(\cdot\right)$ are spline basis functions (linear or log-linear)

Dynamic Panel: Model

Structure of idiosyncratic errors:

$$\begin{aligned} \varepsilon_{it} &= \varepsilon_{it-1} + \xi_{it} \\ \xi_{it} &= X_i'\lambda + \upsilon_{it} \end{aligned}$$

Now rewrite labor demand as

$$\begin{aligned} \ln(L_{it}) - \ln(L_{it-1}) &= -\frac{1-\alpha}{\beta}(\ln w_{it} - \ln w_{it-1}) + \frac{\delta_t - \delta_{t-1}}{\beta} D_i \\ &\quad + \sum_{k=1}^{3} \frac{\theta_k}{\beta}\left[g_k\left(\frac{L_{it-1}}{R_i}\right) - g_k\left(\frac{L_{it-2}}{R_i}\right)\right] \\ &\quad + X_i'\widetilde{\lambda} + \widetilde{\gamma}_t - \widetilde{\gamma}_{t-1} + \widetilde{\upsilon}_{it} \end{aligned}$$

Here county FE removed by differencing
Tildes dnote renormalizing by $\frac{1}{\beta}$
Now $\frac{\delta_t - \delta_{t-1}}{\beta}$ are change in direct effect of TVA between decade
Spline coefficients $\frac{\theta_k}{\beta}$ measure indirect effects of program

Dynamic Panel: Model

$$
\begin{aligned}
\ln\left(L_{it}\right)-\ln\left(L_{it-1}\right) &= -\frac{1-\alpha}{\beta}\left(\ln w_{it}-\ln w_{it-1}\right)+\frac{\delta_t-\delta_{t-1}}{\beta}D_i \\
&+\sum_{k=1}^{3}\frac{\theta_k}{\beta}\left[g_k\left(\frac{L_{it-1}}{R_i}\right)-g_k\left(\frac{L_{it-2}}{R_i}\right)\right] \\
&+X_i'\widetilde{\lambda}+\widetilde{\gamma}_t-\widetilde{\gamma}_{t-1}+\widetilde{\upsilon}_{it}
\end{aligned}
$$

Identification difficult. One problem is agglomeration function

Ideal is randomly assign plants, measure shape of agglomeration function.

Approach: instrument density with 20-year lagged density.

Instruments, $k \in \{1,2,3\}$:

$$
Z_{it}^{(k)} \equiv g_k\left(\frac{L_{it-2}}{R_i}\right)-g_k\left(\frac{L_{it-3}}{R_i}\right)
$$

Identifying assumption:

$$
\mathbb{E}\left[v_{it}Z_{it}^{(k)}\right]
$$

Dynamic Panel: Model

$$
\begin{aligned}
\ln\left(L_{it}\right) - \ln\left(L_{it-1}\right) &= -\frac{1-\alpha}{\beta}\left(\ln w_{it} - \ln w_{it-1}\right) + \frac{\delta_t - \delta_{t-1}}{\beta} D_i \\
&\quad + \sum_{k=1}^{3} \frac{\theta_k}{\beta}\left[g_k\left(\frac{L_{it-1}}{R_i}\right) - g_k\left(\frac{L_{it-2}}{R_i}\right)\right] \\
&\quad + X_i' \widetilde{\lambda} + \widetilde{\gamma}_t - \widetilde{\gamma}_{t-1} + \widetilde{\upsilon}_{it}
\end{aligned}
$$

Identification difficult. Other problem is wages
Solution: calibrate labor demand elasticity $-\frac{1-\alpha}{\beta}$ from Hammermesh literature review.

Dynamic Panel: Results

Table V: Structural Estimates of Agglomeration Function (linear basis)

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
Change in Manufacturing Density Spline Components:						
	0.066	0.060	0.060	0.097	0.084	0.082
Low	(0.010)	(0.010)	(0.010)	(0.037)	(0.037)	(0.037)
				[129.06]	[122.25]	[121.07]
	0.021	0.022	0.022	0.042	0.043	0.042
Medium	(0.0045)	(0.0044)	(0.0044)	(0.011)	(0.011)	(0.011)
				[116.87]	[114.87]	[116.66]
	-0.000075	0.000031	-0.00011	0.0019	0.0021	0.0018
High	(0.00074)	(0.00075)	(0.00071)	(0.0020)	(0.0020)	(0.0020)
				[41.82]	[40.96]	[32.04]
Log Manufacturing Wages	**-1.5**	**-1.5**	**-1.5**	**-1.5**	**-1.5**	**-1.5**
	0.0033	0.0081	0.0026	-0.0052	0.0012	-0.0043
TVA	(0.015)	(0.015)	(0.015)	(0.017)	(0.016)	(0.016)
Regional Trends	no	no	yes	no	no	yes
1940 Manufacturing Density	no	yes	yes	no	yes	yes
Decade Effects	yes	yes	yes	yes	yes	yes
Controls for 1920 and 1930 characteristics	yes	yes	yes	yes	yes	yes
P-value equal slopes	2.7e-11	1.1e-10	9.5e-11	7.1e-04	6.0e-04	6.4e-04
P-value slopes equal zero	1.5e-10	6.0e-10	4.8e-10	.0022	.0015	.0019
N	6057	6057	6057	5952	5952	5952

Dynamic Panel: Results

Table VI: Structural Estimates of Agglomeration Function (log basis)

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	2SLS	2SLS	2SLS
Change in Log Manufacturing Density Spline Components:						
	0.173	0.147	0.146	0.443	0.400	0.396
Low	(0.037)	(0.037)	(0.037)	(0.102)	(0.108)	(0.107)
				[177.17]	[159.14]	[157.20]
	0.221	0.227	0.226	0.456	0.440	0.438
Medium	(0.045)	(0.044)	(0.045)	(0.124)	(0.123)	(0.124)
				[106.74]	[109.55]	[110.13]
	0.143	0.151	0.141	0.466	0.467	0.453
High	(0.051)	(0.050)	(0.050)	(0.150)	(0.150)	(0.151)
				[206.66]	[204.69]	[200.36]
Log Manufacturing Wages	**-1.5**	**-1.5**	**-1.5**	**-1.5**	**-1.5**	**-1.5**
	0.007	0.012	0.008	-0.003	0.002	-0.002
TVA	(0.014)	(0.014)	(0.014)	(0.012)	(0.013)	(0.012)
Regional Trends	no	no	yes	no	no	yes
1940 Manufacturing Density	no	yes	yes	no	yes	yes
Decade Effects	yes	yes	yes	yes	yes	yes
Controls for 1920 and 1930 characteristics	yes	yes	yes	yes	yes	yes
P-value equal slopes	.2483	.1298	.1038	.9545	.669	.7171
P-value slopes equal zero	1.9e-07	5.1e-07	6.9e-07	1.5e-04	7.4e-04	.001
N	6057	6057	6057	5935	5935	5935

Dynamic Panel: Results

Table VII: Direct Effects of TVA on Labor Demand, by Period

	(1) 1940-1960	(2) 1960-1980	(3) 1980-2000
Spline in Levels	0.225***	-0.011	-0.004
	(0.070)	(0.041)	(0.039)
Spline in Logs	0.185**	-0.036	0.033
	(0.081)	(0.033)	(0.035)
Controls for 1920 and 1930 characteristics	yes	yes	yes
N	1587	1498	1533

Dynamic Panel: Benefit-Cost Analysis

Elasticity of utility w.r.t. TVA investments:

$$\frac{d\overline{u}}{d\delta} = \frac{1}{1-\alpha} \frac{\sum_i \frac{D_i L_i}{\beta - \sigma_i}}{\sum_i \frac{L_i}{\beta - \sigma_i}}$$

With constant elasticity ($\sigma_i = \sigma$), this simplifies to

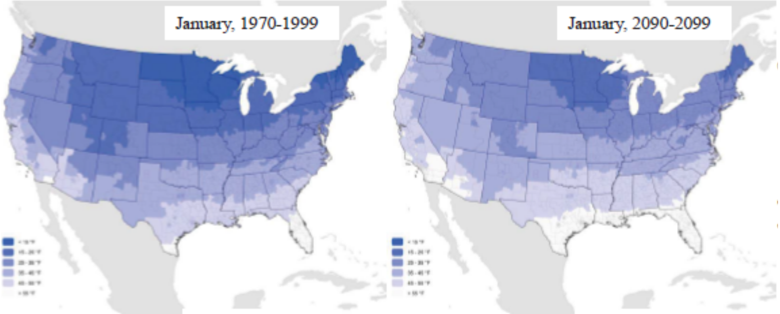
$$\frac{d\overline{u}}{d\delta} = \frac{1}{1-\alpha} \frac{\sum_i D_i L_i}{\sum_i L_i}$$

Substituting with discount rate 3T, benefits were \$24 to \$37 billion (depending on labor supply elasticity) and costs \$17.3 billion.

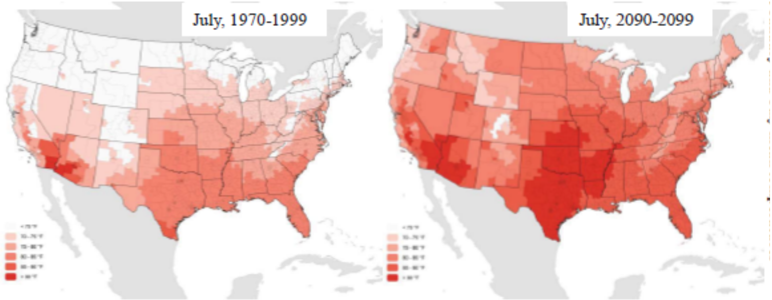
Climate Applications of Roback/Rosen

- Climate change predicted to harm agriculture
- Until 1990s, most evidence used agronomy optimization models from farm research
- AER paper argued for measuring cross-sectional relationship of land values and climate (Mendelsohn, Nordhaus, and Shaw 1994)
- Huge subsequent literature builds on this to analyze how climate change affects land values, inferring effects on agriculture

Albouy et al. “Climate Amenities”



Albouy et al. "Climate Amenities"



Albouy et al. "Climate Amenities"

- Some existing papers estimate willingness to pay (WTP), this builds on them in three ways
  - Account for local expenditures besides housing, and net out federal taxes from local wage differences
  - Use full temperature distribution, not just seasonal or monthly means
    - Useful to learn about extremes
  - Analyze unobserved heterogeneity in preferences via Bajari and Benkard (2005)

Albouy et al. Model (Environment?)

- j locations
- i mobile, price-taking households
- x traded numeraire good; y non-traded home good
- p_j price of local good y (determines cost of living)
- w_j^i Households' wage income, varies by location j
- R_i: rental income from households portfolios of land and capital
- $m_j^i = R^i + w_j^i$: gross household income
- $\tau\left(m_j^i\right)$: federal taxes, rebated lump sum
- Z_j: K-dimensional amenity vector, including climate
- ξ_j : scalar characteristic that households observe but we don't.

Albouy et al. Model

- Household problem: choose location j to maximize indirect utility $u_j^i = V^i\left(p_j, w_j^i, Z_j, \xi_j\right)$.
  - Heterogeneous preferences over (Z, ξ)
- Firm problem: competitive input and output markets, zero profits
  - Offer a higher wage in more productive locations
  - $\varphi^i w_j$ is wage of household i in location j
  - φ^i: household-specific skill level

Albouy et al. Model

- Equilibrium: $p\left(Z_j, \xi_j\right)$ and $w\left(Z_j, \xi_j\right)$
  - Depend on households' demand for amenities, firms' location decisions, and land supply
  - Heterogeneous preferences over (Z, ξ)
- FOC$[Z_k]$ from household utility maximization

$$\underbrace{\frac{1}{m_j^i \lambda^i} \frac{\partial V^i\left(p_j, w_j, Z_j, \xi_j\right)}{\partial Z_k}}_{\text{MWTP for attribute } k} = \underbrace{\frac{p_j y_j^i}{m_j^i} \frac{\partial \ln p\left(Z_j, \xi_j\right)}{\partial Z_k}}_{\text{Log change in cost of living}} - \underbrace{\frac{w_j^i\left(1-\tau'\right)}{m_j^i} \frac{\partial \ln w\left(Z_j, \xi_j\right)}{\partial Z_k}}_{\text{Log change in wages}} \tag{1}$$

λ^i marginal utility of income
τ' average marginal tax rate on labor income

Albouy et al. Model

- Empirical version:

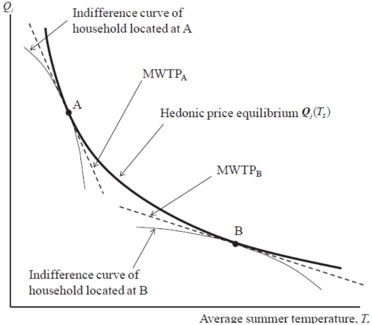
$$\hat{Q}^j = s_y \hat{p}_j - \left(1 - \tau'\right) s_w \hat{w}_j$$

$$= 0.33\hat{p}_j - 0.50\hat{w}_j$$

- s_y: share of income spent on local goods
- s_w: share of income from wages
- hats: differentials relative to US income-weighted average.
- values 0.33, 0.50 from Albouy (2011)

Albouy et al. Model

Figure 1: Illustrative hedonic price function with demand-side equilibrium FOC's satisfied



Intuition: cooler weather is "paid for" via higher housing prices or lower wages.
(In this graph, people dislike heat.)
Slope of $\hat{Q}(T_s)$ at a location is hedonic price for a marginal increase in temperature at that location.
Also equals MWTP for households at that location (see example households A and B).

Albouy et al. Data

- IPUMS 2000 5-percent microdata sample
- Smallest geography: Public Use Microdata Area (PUMA), population 100,000-300,000
- About 2,000 of these (NB: US has 3,000 counties)
- Wage differentials

$$\ln w_{ij} = X_{ij}^{w} \beta^{w} + \mu_j^{w} + \varepsilon_{ij}^{w}$$

  - Full-time workers aged 25-55
  - Regress log-wage of worker i in location j on PUMA dummies μ_j^w, controls X_{ij}^w (interacted with gender) for education, experience, race, occupation, industry, veteran, marital, immigration.
  - μ_j^w: regression estimate of PUMA wage differential $\hat{w}_j$
  - Interpretation: causal effect of a PUMA's characteristics on a worker's wage
  - Threats to internal validity

Wage Differentials

- Wage differentials

$$\ln w_{ij} = X_{ij}^{w} \beta^{w} + \mu_j^{w} + \varepsilon_{ij}^{w}$$

  - Full-time workers aged 25-55
  - Regress log-wage of worker i in location j on PUMA dummies μ_j^w, controls X_{ij}^w (interacted with gender) for education, experience, race, occupation, industry, veteran, marital, immigration.
  - μ_j^w: regression estimate of PUMA wage differential $\hat{w}_j$
  - Interpretation: causal effect of a PUMA's characteristics on a worker's wage
  - Threats to internal validity
    - Workers must not sort across places of work according to unobserved skills
    - Good assumption?
    - Ways to improve on this?

Housing Differentials

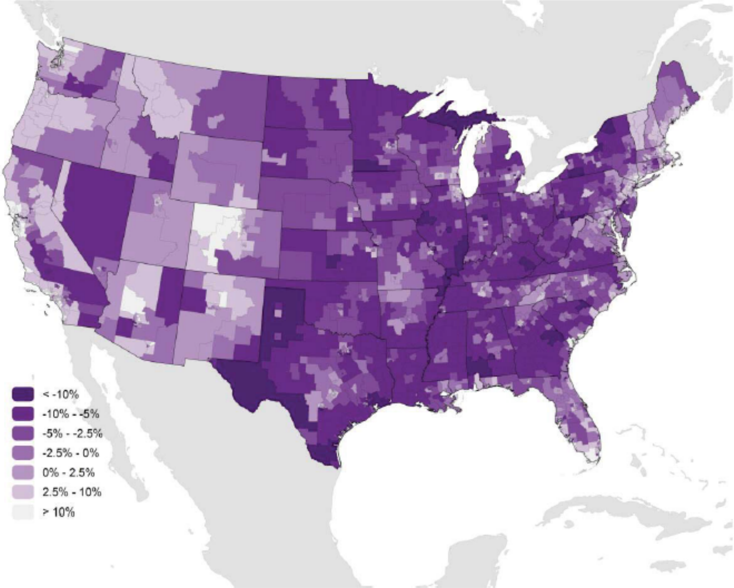
- Housing differentials

$$\ln p_{ij} = Y_{ij}\beta^p + \mu_j^p + \varepsilon_{ij}^p$$

  - Covariates Y_{ij} (interacted with renter status): size, rooms, acreage, commercial use, kitchen, plumbing facilities, building type, building age, residents per room
  - Covert housing values to rents (discount rate 7.85 percent)
  - Add utility costs (electricity, gas, oil)
  - μ_j^p: estimate of PUMA housing cost differential $\hat{p}_j$
  - Identifying assumption: unobserved housing quality doesn't vary systematically across locations
  - Notes:
    - They want to recover the fixed effect, not a parameter
    - Don't show the regression equation
    - Would panel data help?
    - Should they include utility costs?

Housing Differentials

Figure 2: Quality of life (QOL) differentials in 2000



Housing Differentials

Table 1: Descriptive statistics for primary dataset.
Sample consists of 2057 Public Use Microdata Areas (PUMAs)

	Mean	Std. Dev	10th percentile	90th percentile
Temperature data, 1970-1999 average				
Average annual heating degree days (1000s)	4.384	2.204	1.326	7.009
Average annual cooling degree days (1000s)	1.290	0.929	0.374	2.762
Temperature data, 2090-2099 projected (CCSM A2)				
Projected avg. annual heating degree days (1000s)	2.821	1.698	0.425	4.845
Projected avg. annual cooling degree days (1000s)	2.747	1.234	1.447	4.719
Other climate data, 1970-1999 average				
Average annual precipitation (inches)	39.26	14.09	16.25	53.85
Average annual relative humidity (%)	63.58	8.10	53.31	70.52
Average annual sunshine (% of available daylight)	60.17	8.64	49.74	73.08
Other climate data, 2090-2099 projected (CCSM A2)				
Projected avg. annual precipitation (inches)	41.60	15.56	14.68	56.91
Projected avg. annual relative humidity (%)	62.41	8.94	50.64	69.44
Projected avg. annual sunshine (% of available daylight)	61.55	8.63	51.29	72.81

Housing Differentials

Geographic data				
Distance from centroid of puma to ocean (miles)	250.1	272.1	4.3	729.2
Distance from centroid of puma to Great Lake (miles)	763.2	715.4	54.0	2128.4
Average land slope (degrees)	1.677	2.131	0.191	4.270
Demographic data (2000 census)				
Weighted population density (people per sq. mile)	5,466	11,997	360	9,981
Percent high school graduates	80.1%	10.0%	67.2%	90.9%
Percent of population with bachelors degree	24.1%	12.4%	11.3%	41.0%
Percent of population with graduate degree	8.7%	5.6%	3.7%	20.3%
Average age	48.7	2.6	45.3	51.8
Percent hispanic	8.9%	13.6%	0.6%	81.6%
Percent black	12.3%	17.6%	0.5%	96.7%
Population (1000s)	135.9	32.9	103.8	182.2
Quality of life differential (in logs)	0.034	0.071	-0.043	0.128

Data

Geography: 2,057 PUMAs
Climate: 1970-1999 mean temperatures, precipitation, humidity, sunshine
Climate forecasts: business-as-usual, Community Climate System Model (CCSM)
Other controls: distance to ocean and great lakes; mean slope; population density, educational attainment; age; racial-ethnic composition

Homogenous Willingness to Pay

$$\hat{Q}_j = \beta X_j + \gamma D_j + \xi_j$$

$\hat{Q}$: quality of life differential
X_j: climate variables
D_j: control variables

Homogenous Willingness to Pay

Temperature spline:

$$f\left(t\right)=\sum_{s=1}^{S}\beta_{s}S_{s}\left(t\right)$$

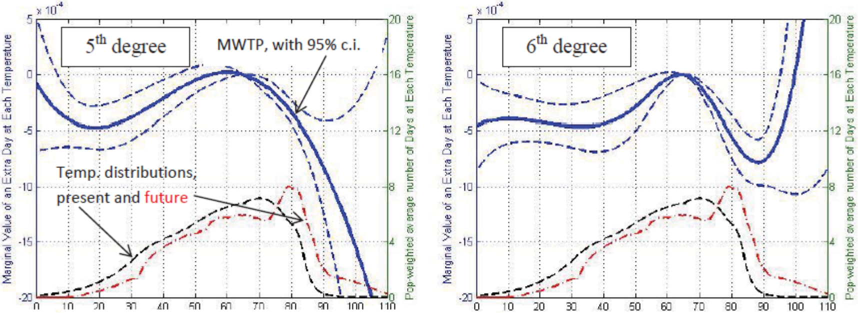
Substituting,

$$\begin{aligned}\hat{Q}_{j} &= \sum_{t}N_{jt}f\left(t\right)+\beta_{r}Rain_{j}+\beta_{h}Humd_{j}+\beta_{s}Sun_{j}+\gamma D_{j}+\gamma_{j}\\ &= \sum_{s=1}^{S}\beta_{s}\left(\sum_{t}N_{jt}S_{s}\left(t\right)\right)+\beta_{r}Rain_{j}+\beta_{h}Humd_{j}+\beta_{s}Sun_{j}\\ &\quad +\gamma D_{j}+\xi_{j}\end{aligned}$$

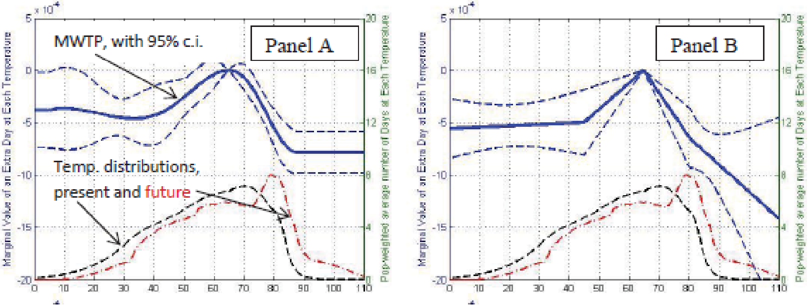
GLS, weighted by PUMA population
Cluster standard errors by state

Homogenous Willingness to Pay

Figure 4: Estimated WTP for temperature, homogenous preference cubic spline models



Homogenous Willingness to Pay



Homogenous Willingness to Pay

Table 2: Estimation results for the linear, homogenous preferences model. Dependent variable is quality of life (QOL) differential as fraction of income

	I	II	III	IV	V	VI
Temperature function	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline
MWTP for a day at 40F (x365)	-0.158 (0.052)	-0.184 (0.051)	-0.179 (0.061)	-0.193 (0.056)	-0.177 (0.059)	-0.216 (0.053)
MWTP for a day at 80F (x365)	-0.203 (0.033)	-0.236 (0.053)	-0.217 (0.043)	-0.297 (0.087)	-0.188 (0.055)	-0.227 (0.075)
Average annual precipitation (inches, x10)	0.008 (0.002)	0.008 (0.002)	0.003 (0.003)	0.003 (0.003)	0.008 (0.003)	0.008 (0.003)
Average annual relative humidity (percent)	0.013 (0.051)	-0.001 (0.045)	0.023 (0.077)	0.025 (0.059)	-0.019 (0.051)	-0.031 (0.052)
Average annual sunshine (fraction of all daytime hours)	0.198 (0.054)	0.200 (0.052)	0.176 (0.078)	0.179 (0.068)	0.104 (0.079)	0.118 (0.079)
Average land slope (degrees)	0.005 (0.001)	0.004 (0.001)	0.007 (0.002)	0.007 (0.002)	0.004 (0.001)	0.004 (0.001)
Inverse distance to ocean (miles)	0.233 (0.032)	0.231 (0.032)	0.341 (0.044)	0.340 (0.045)	0.225 (0.036)	0.218 (0.036)

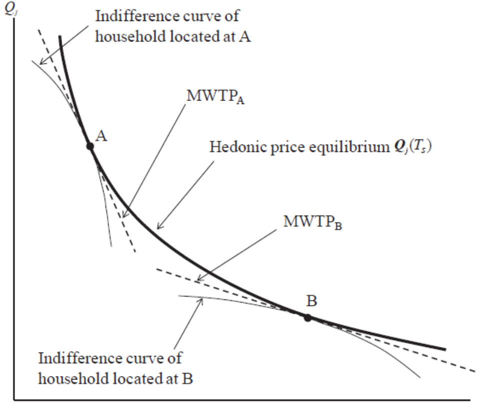
Homogenous Willingness to Pay

Inverse distance to Great Lakes (miles)	0.144 (0.110)	0.154 (0.110)	0.004 (0.191)	0.034 (0.183)	0.080 (0.112)	0.089 (0.111)
Squared inverse distance to ocean (miles)	-0.219 (0.029)	-0.217 (0.030)	-0.296 (0.061)	-0.294 (0.059)	-0.224 (0.031)	-0.219 (0.032)
Squared inverse distance to Great Lakes (miles)	-0.145 (0.122)	-0.156 (0.121)	0.089 (0.209)	0.054 (0.197)	-0.092 (0.122)	-0.101 (0.122)
Log of weighted population density	0.004 (0.002)	0.004 (0.002)	-	-	0.002 (0.002)	0.002 (0.002)
Fraction high school graduates	0.109 (0.029)	0.107 (0.030)	-	-	0.098 (0.027)	0.095 (0.027)
Fraction bachelors degrees	0.278 (0.047)	0.272 (0.049)	-	-	0.284 (0.048)	0.286 (0.049)
Fraction graduate degrees	0.039 (0.078)	0.048 (0.086)	-	-	0.062 (0.074)	0.060 (0.076)
Average age (x100)	0.152 (0.050)	0.148 (0.049)	-	-	0.154 (0.041)	0.147 (0.041)
Fraction hispanic	0.048 (0.028)	0.043 (0.032)	-	-	0.057 (0.022)	0.055 (0.024)
Fraction black	-0.072 (0.007)	-0.074 (0.007)	-	-	-0.064 (0.008)	-0.066 (0.007)
State fixed effects	N	N	N	N	Y	Y
Number of observations	2057	2057	2057	2057	2057	2057
R^2	0.782	0.780	0.366	0.364	0.822	0.822

Heterogeneous Willingness to Pay

Step 1: flexibly estimate hedonic price function
Obtain local price for climate at each location
Step 2: household's local MWTP given by FOC
Downside: only identifies MWTP for each household from its chosen climate, then linearly extrapolate

Figure 1: Illustrative hedonic price function with demand-side equilibrium FOC's satisfied



Heterogeneous Willingness to Pay

Approach: estimate $\hat{Q}\left(Z_j, \xi_j\right)$ flexibly, then use local linear regression to recover preferences.

$$\hat{Q}_j = \sum_k \beta_k^{j*} Z_{jk} + \xi_j$$

β_k^{j*}: implicit prices for climate amenities. Estimate via weighted least squares

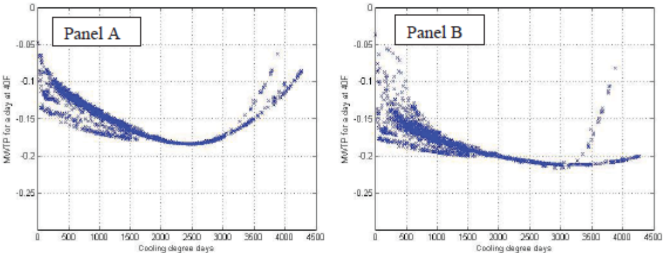
$$\begin{aligned}
\beta^{j*} &= \arg\min_{\beta} \left(Q - Z\beta\right)' W \left(Q - Z\beta\right) \\
Q &= \left[Q_j\right] \\
Z &= \left[Z_j\right] \\
W &= diag\left[K_h\left(Z_j - Z_{j^*}\right)\right]
\end{aligned}$$

Heterogeneous Willingness to Pay

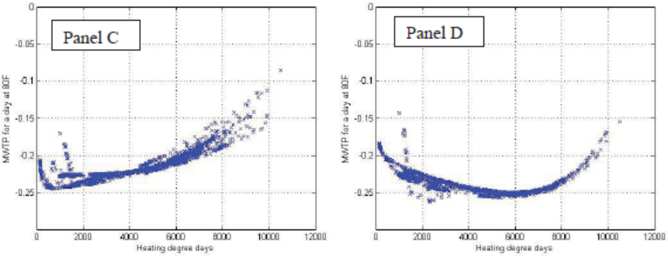
Kernel weights defined so locations with similar characteristics have most influence:

$$\begin{aligned} K(Z) &= \prod_k N\left(\frac{Z_k}{\hat{\sigma}_k}\right) \\ K_h(Z) &= K\left(\frac{Z}{h}\right)\frac{1}{h} \end{aligned}$$

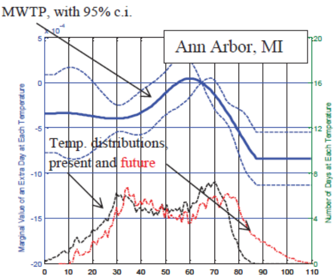
Heterogeneous Willingness to Pay

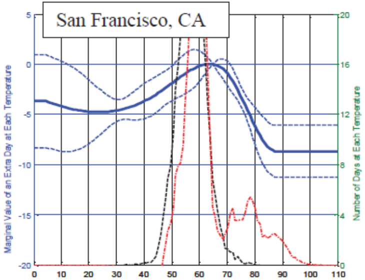


Heterogeneous Willingness to Pay



Heterogeneous Willingness to Pay





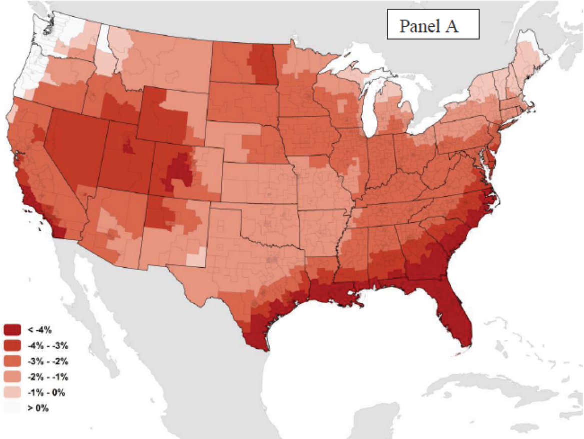
Heterogeneous Willingness to Pay

Table 3: Estimation results for the heterogeneous preferences model. Dependent variable is quality of life (QOL) differential as fraction of income

	I	II	III	IV	V	VI
Temperature function	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline
Average MWTP for a day at 40F (x365)	-0.145 (0.047)	-0.179 (0.045)	-0.169 (0.057)	-0.201 (0.051)	-0.175 (0.051)	-0.215 (0.044)
Average MWTP for a day at 80F (x365)	-0.215 (0.023)	-0.241 (0.045)	-0.283 (0.045)	-0.357 (0.063)	-0.205 (0.046)	-0.236 (0.060)
Geog & other weather controls	Y	Y	Y	Y	Y	Y
Demographic controls	Y	Y	N	N	Y	Y
State fixed effects	N	N	N	N	Y	Y

Heterogeneous Willingness to Pay

Figure 8: Predicted changes in QOL as percent of income under the A2 scenario for 2090-2099, with the heterogeneous preference model MWTP estimates



Heterogeneous Willingness to Pay

Table 5: Projected U.S. average amenity value impacts from climate change for 2090-2100 using the preference estimates from the heterogeneous preference models.

	I	II	III	IV	V	VI
Temperature function	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline
A2 scenario change in welfare as percent of income	-2.83% (0.30%)	-3.84% (0.76%)	-3.34% (0.39%)	-2.49% (0.97%)	-1.24% (0.82%)	-0.87% (1.31%)
A1FI scenario change in welfare as percent of income	-3.37% (0.36%)	-4.74% (1.01%)	-3.88% (0.48%)	-2.85% (1.33%)	-1.69% (0.98%)	-1.25% (1.67%)
Geog & other weather controls	Y	Y	Y	Y	Y	Y
Demographic controls	Y	Y	N	N	Y	Y
State fixed effects	N	N	N	N	Y	Y

Notes : Parenthetical values indicate standard errors wild cluster bootstrapped on state using 200 draws.

Summary

- Research: what is MWTP for climate amenities for residential US?
- Methodology: cross-sectional hedonic model, Bajari-Benkard to allow heterogeneous preferences
- Conclusions:
  - Americans prefer a temperature near 65 F (agrees with degree-day models)
  - On margin, households prefer avoiding excess heat than excess cold
  - Climate change cost in 2100 (8.3 F) = 1 to 3 percent of welfare

Comments

- Cross-sectional identification?
- Challenges for extrapolating to 2100:
  - Adaptation: preferences, technologies change
  - Linear MWTP extrapolation strong
- Ambiguous what this covers
  - Agricultural productivity?
  - Extreme weather risks?
- No uncertainty here

Open Research Questions

- Recovering primitives: nonlinear indifference curves as function of demographics
- Other amenities: crime? tax and healthcare policy?
- Incorporate adaptation?
- China? India? US a tiny share of global population
- How incorporate these numbers into integrated assessment models?