The Impact of Local Labor Market Conditions on Migration: Evidence from the Bakken Oil Boom

Mallory C. Vachon
Center for Energy Studies
Louisiana State University
December 1, 2015

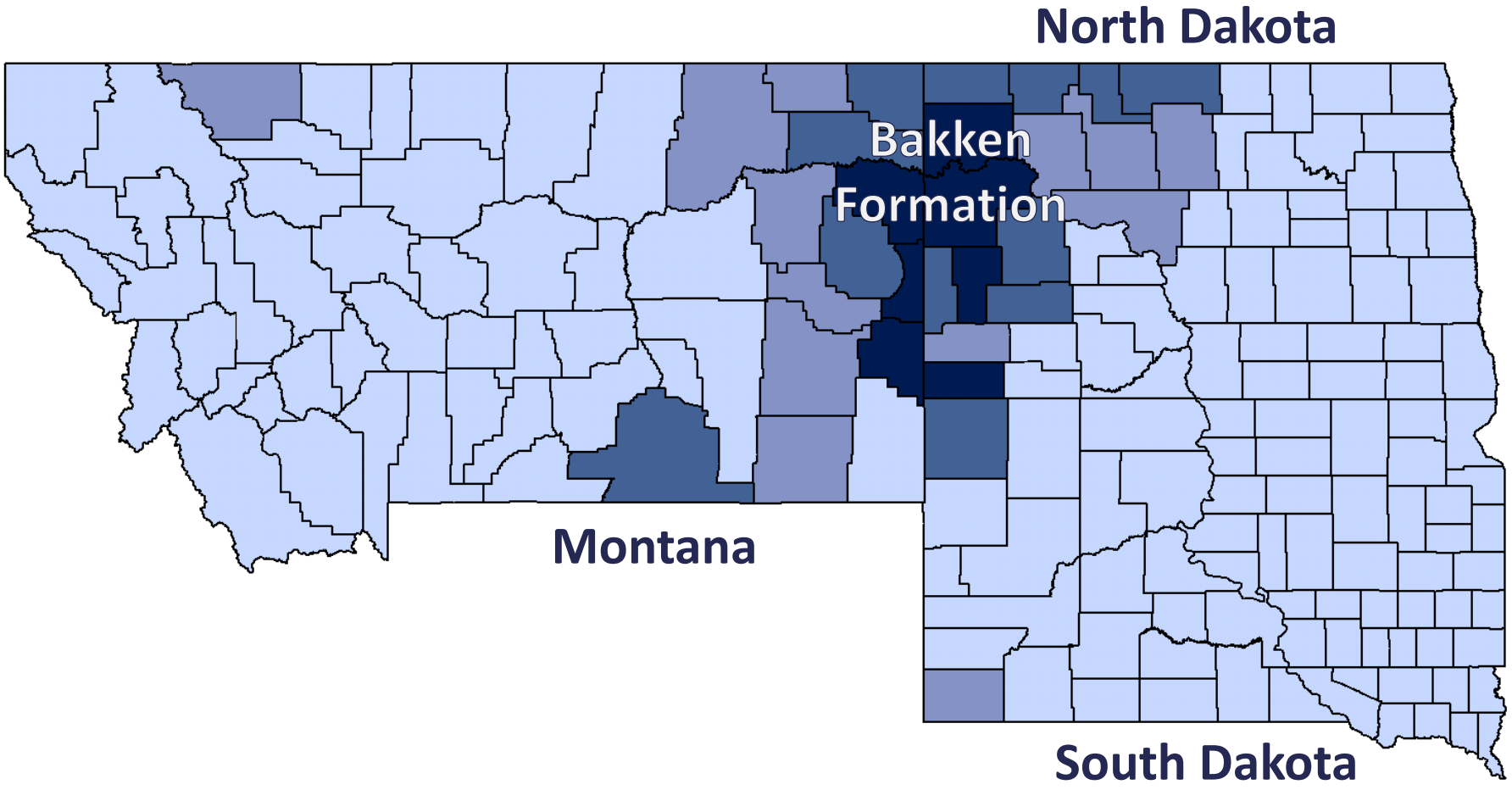
Motivation

- Migration is a primary mechanism of **regional labor market adjustment**
- Existing literature provides **few causal estimates** of the relationship between local labor market conditions and migration
- In this paper, I provide causal estimates of the impact of exogenous change in earnings on net migration

Background

- **Oil boom** in Bakken Formation of the Williston Basin impacting Montana (MT), North Dakota (ND), and South Dakota (SD)
- Boom led to an **exogenous labor demand shock** that increased earnings
- Production increased from 50 million barrels in 2000 to 250 million barrels in 2010
- Part of a larger boom in oil and natural gas production in the United States

Oil Reserves: MT, ND, SD



Empirical Challenge

- Fundamental challenge in estimating the relationship between local economic conditions and migration
- At the local labor market level, earnings, employment, and migration are **jointly determined**
- Implement an instrumental variable (IV) strategy that relies on three sources of variation
  - Oil reserves
  - Oil prices
  - Technology

Data

- Internal Revenue Service (IRS)
  - County-level migration
  - Wage and salary earnings
- Energy Information Administration (EIA)
  - County-level oil reserves
  - West Texas Intermediate (WTI) crude oil prices

Summary of Results

- Earnings growth in oil counties significantly **increases** net migration
- Net migration rate in North Dakota oil counties increased by 2.6 percentage points
- Net migration rate in the three-state region increased by 3.2 percentage points

Summary of Results

- Economically sizeable
  - Pre-boom net migration rate was -1.5 percent (out-migration)
  - Impact of boom is +2.6 percentage points
  - Post-boom net migration rate is 1.1 percent (in-migration)
  - The boom has transformed these counties from population-losers to population-gainers, making them among the fastest-growing counties in the country

Outline

- Previous Literature
- Basic Theory
- Econometric Methodology, Data, and Identification Strategy
- Results and Extensions
- Summary

Previous Literature

- Local labor markets literature has examined the impact of local labor supply and demand shocks on labor market outcomes, including earnings, employment, and migration
  - Bartik (1991); Blanchard and Katz (1992); Moretti (2011); Topel (1986)

Previous Literature

- Growing empirical literature examines the impact of natural resources on local economic conditions
- Carrington (1996) examines the impact of Trans-Alaska Pipeline construction on earnings and employment

Previous Literature

- Black, McKinnish, and Sanders (2005) examine the impact of earnings growth on local labor market conditions
- Focus on counties in the four-state region of KY, OH, PA, and WV
- Some counties naturally endowed with coal, others not
- Value of county coal reserves increased and decreased due to a boom and bust in coal prices during the 1970s and 1980s
- Earnings growth during the boom impacts migration:
  - Reduced out-migration of prime-aged men
  - Increased return-migration of prime-aged men
  - Asymmetric shock: out-migration during bust greater than in-migration during boom

Previous Literature

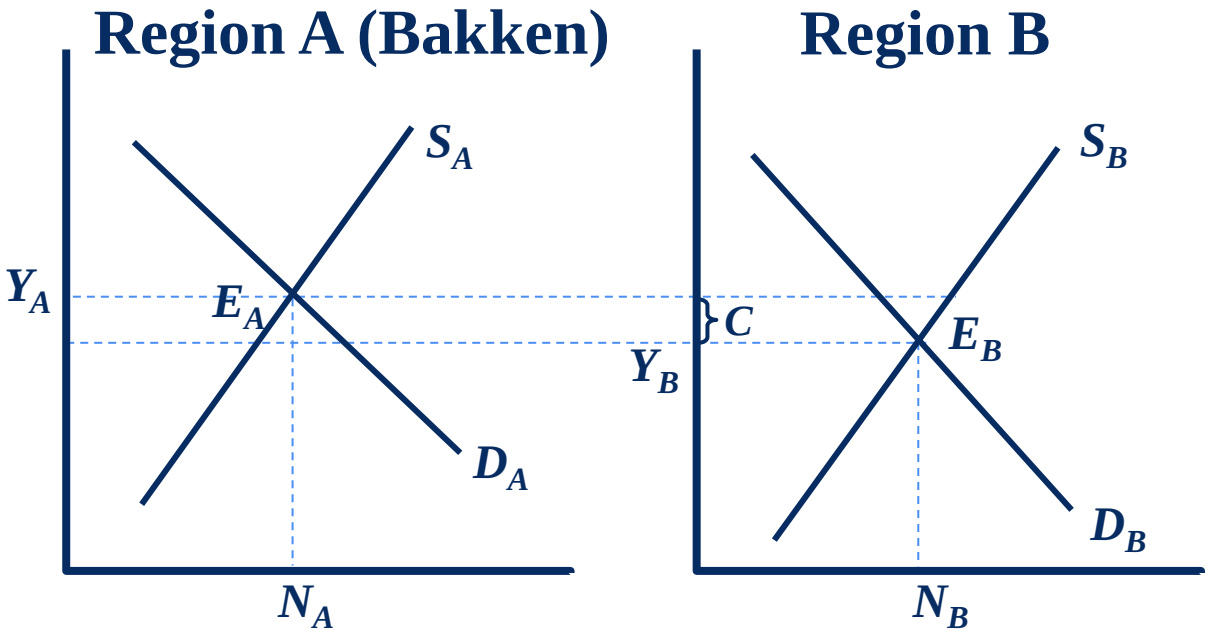
- During coal boom and bust, a negative relationship between the value of labor force participation and Disability Insurance (DI) participation
  - Black, Daniel, and Sanders (2002)
- Resource booms create positive employment spillovers to manufacturing
  - Allcott and Keniston (2014)

Previous Literature

- These papers examine shocks of the 1970s and 1980s
- I examine the recent boom in MT, ND, and SD
  - Technology has changed extractive industries
  - Secular changes in the labor market since the 1970s
- Suggesting that the **responsiveness of migration** to local labor market conditions may be different now than in time periods studied by previous authors

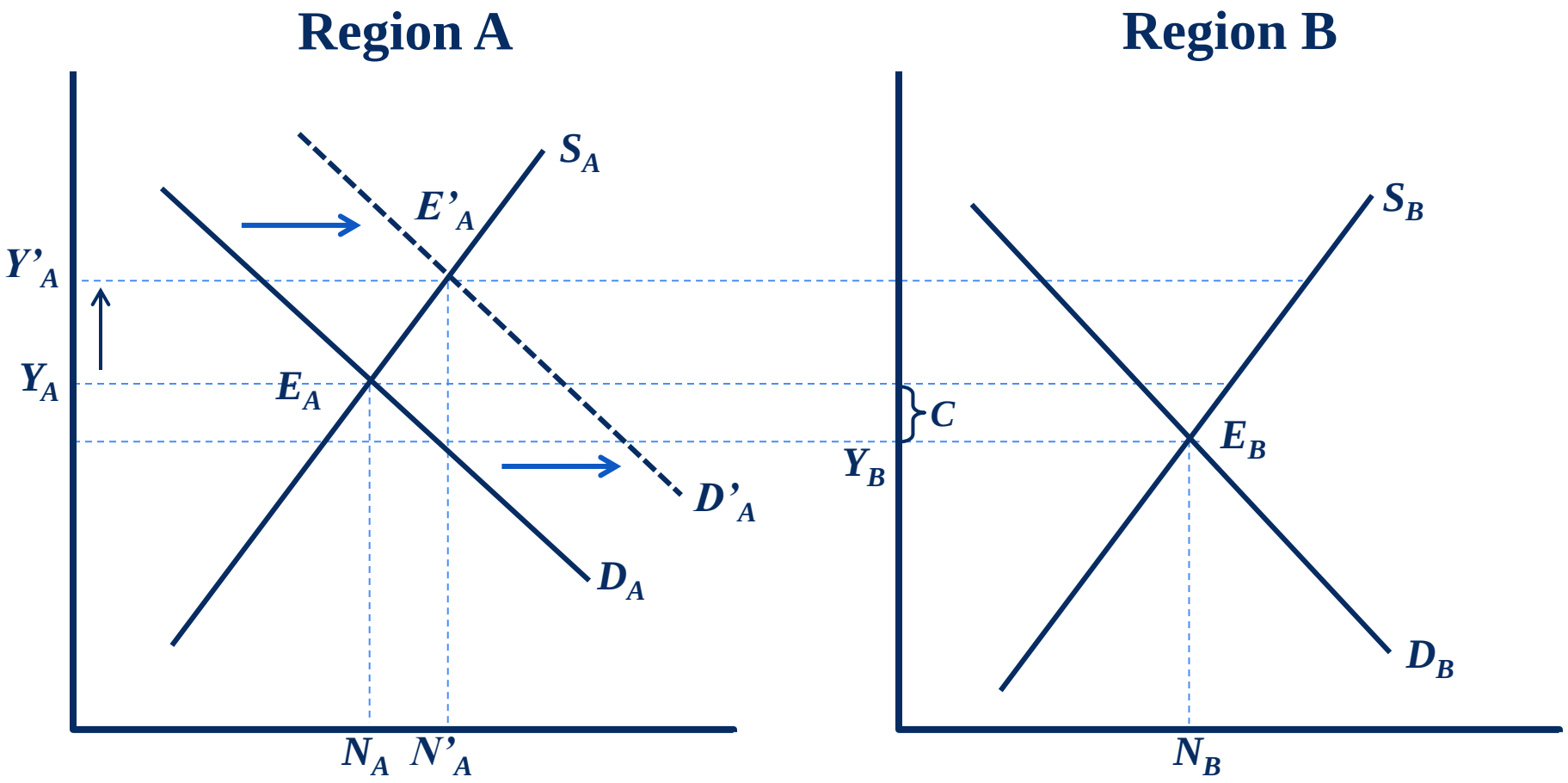
Labor Market Model

- A common feature in the literature is that migration decision is viewed as a **utility maximization** problem
- Utility is typically modeled as a function of **local earnings**, amenities, and the costs of moving
- The migration decision is made based on **earnings differentials**, net of moving costs
- To illustrate, I present a simple two-region model of migration in which labor is the key factor of production
  - Land and capital are assumed to be perfectly elastic in supply

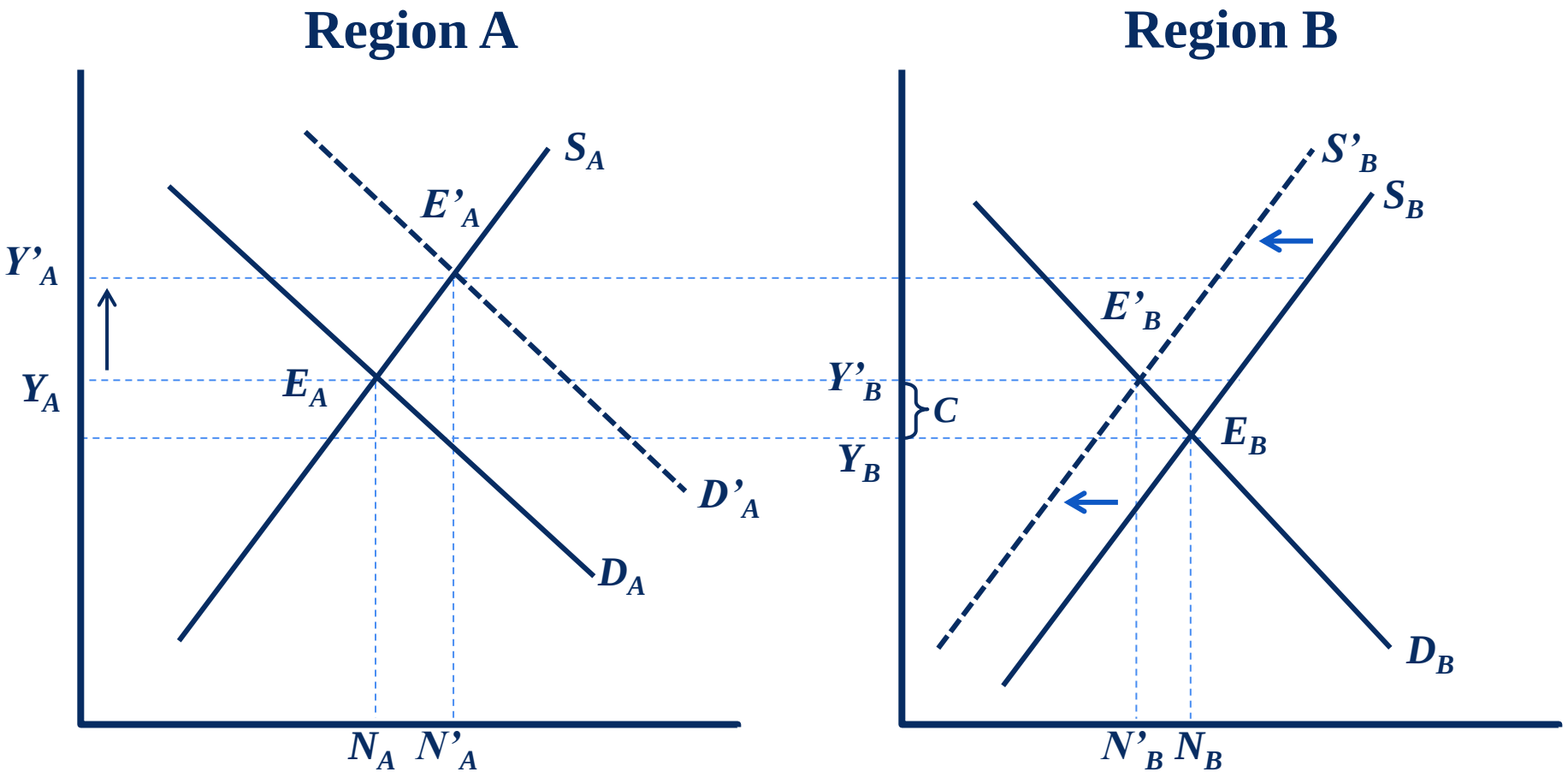


- Labor markets must clear in each region: $D_A = S_A$ and $D_B = S_B$
- Earnings are equal across regions: $Y_A = Y_B + C$

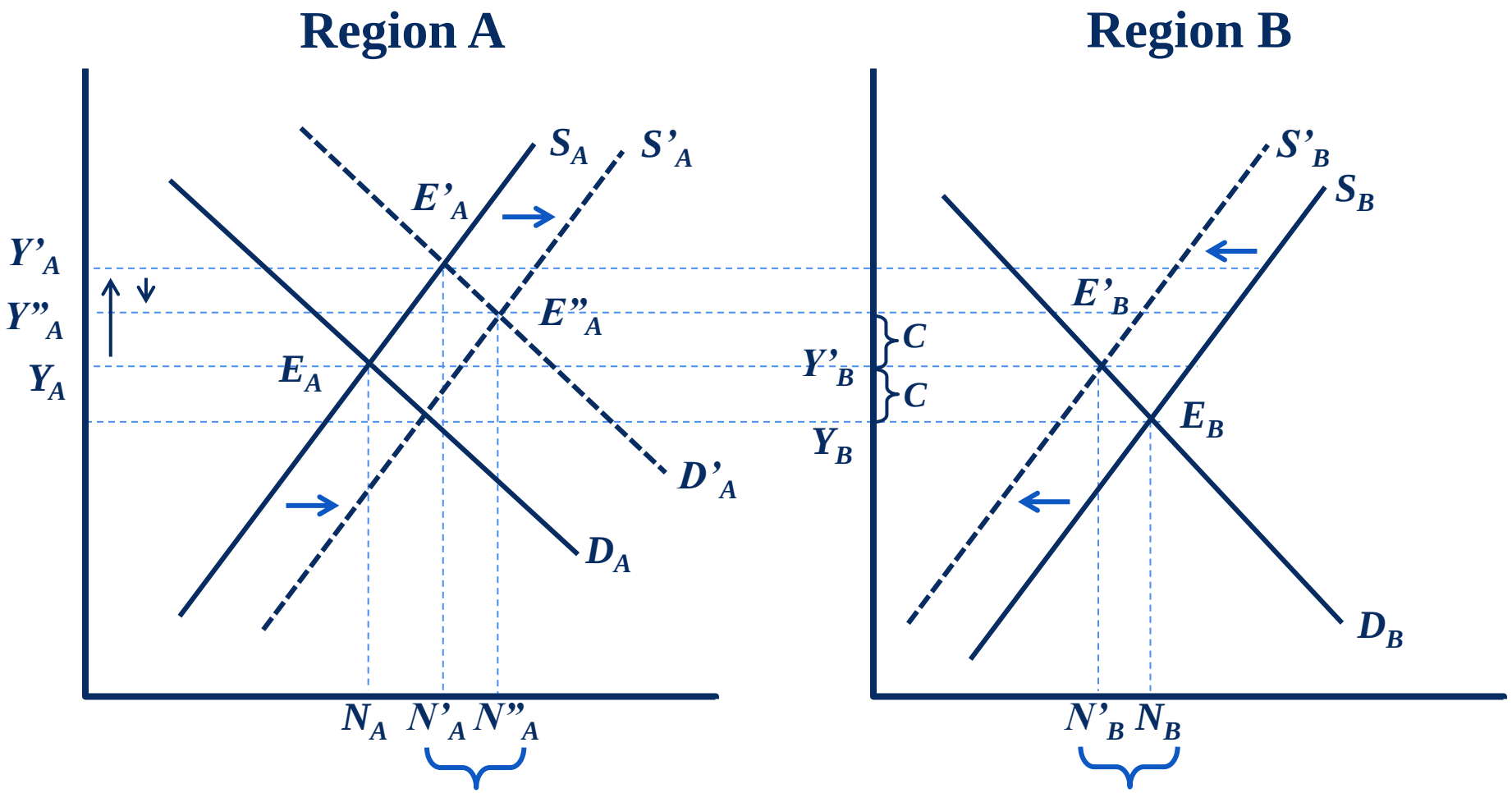
Labor Demand Shock Raises Earnings in Region A

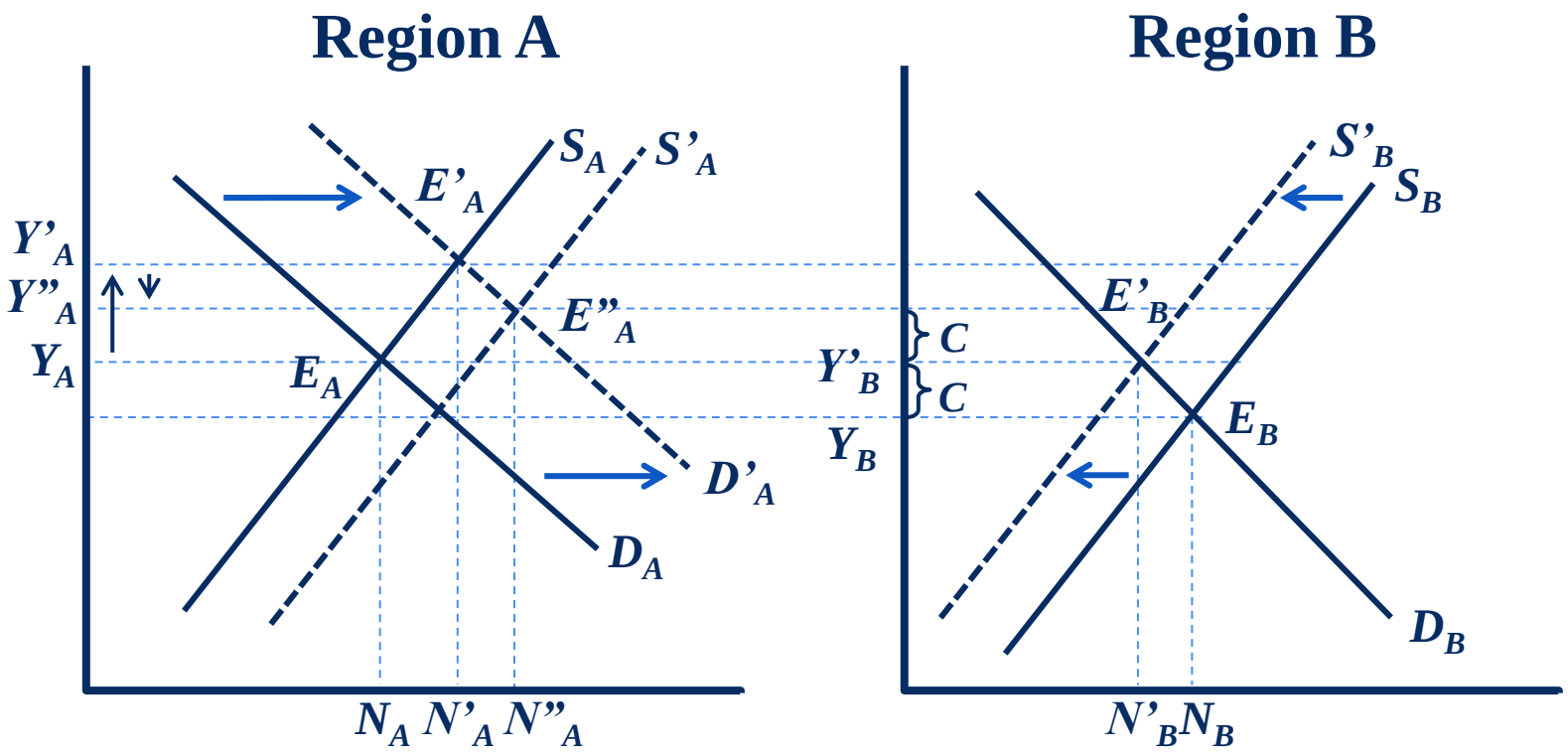


Earnings Differential Induces Out-Migration from Region B



Region B Out-Migration Shifts Region A Labor Supply Curve





- The migration response depends on:
  - The **costs of migration** from Region B to Region A
  - The labor supply and demand **elasticities** in each region

Labor Market Model

- From the model, I derive the following equation that summarizes the relationship between these factors and the migration response:

$$\beta = \eta_A^S - \eta_A^D - \eta_B^S \left(\frac{S_B}{S_A}\right)\left(1 + \frac{C}{Y_B}\right)$$

- β is the semi-elasticity of net migration with respect to earnings: the change in the net migration rate into Region A in response to an increase in earnings in Region A
- As the graphical illustration suggested, β is a function of the labor elasticities and moving costs

Econometric Specification

(1) $m_{ist} = \varphi + \beta ln(y_{ist}) + \pi_i + \tau_t + \phi_{st} + \mu_{ist}$

- County i, state s, year t
- m is the net migration rate
- y is the average household earnings
- π is a county-specific fixed-effect
- τ is a linear time trend
- ϕ is a state-by-year fixed effect
- μ is an iid error

Data - Internal Revenue Service (IRS)

- **Permanent migration** of households measured by tax returns at the **county level**
- I measure the number of households as the number of **returns**
- The IRS defines migration as a year-over-year address change on federal individual income tax returns
- So, a **migrant** is a household that moves to or from county i between years t and $t + 1$

Data - Internal Revenue Service (IRS)

- Net migration flow equals the number of in-migrating households minus the number of out-migrating households
- **Net migration rate** in the time interval t through $t + 1$ equals the net migration flow divided by the number of households in that county in year t

Data - Internal Revenue Service (IRS)

- An important aspect of the Bakken labor market is the presence of temporary migrants
- Workers who live in other states and commute into the Bakken for work
  - For example, work 16 days straight, have 14 days off
- The IRS data do not measure temporary migrants
- This is an important limitation

Data - Internal Revenue Service (IRS)

- County-level **wage and salary earnings** for permanent residents, as measured by federal income tax returns
- Dividing real earnings by the number of returns filed in the county gives the **earnings per household**
- They are expressed in $2010

Table 1 - Sample Means:
Montana, North Dakota, and South Dakota, 1993-2010

Net Migration Rate	-0.008
Total Returns (Households)	5,556
Total Exemptions (Population)	11,799
County Earnings per Return (Thousands of 2010$)	27.03
Number of Counties	175

Two Econometric Challenges

- One challenge is the presence of county fixed effects:

(1) $m_{ist} = \varphi + \beta ln(y_{ist}) + \pi_i + \tau_t + \phi_{st} + \mu_{ist}$

  - Amenities or costs of moving that are location-specific

- The presence of county fixed effects could **bias** my estimates of β

- To account for these, I first-difference (1) :

(2) $\Delta m_{ist} = \delta + \beta \Delta ln(y_{ist}) + \gamma_{st} + u_{ist}$

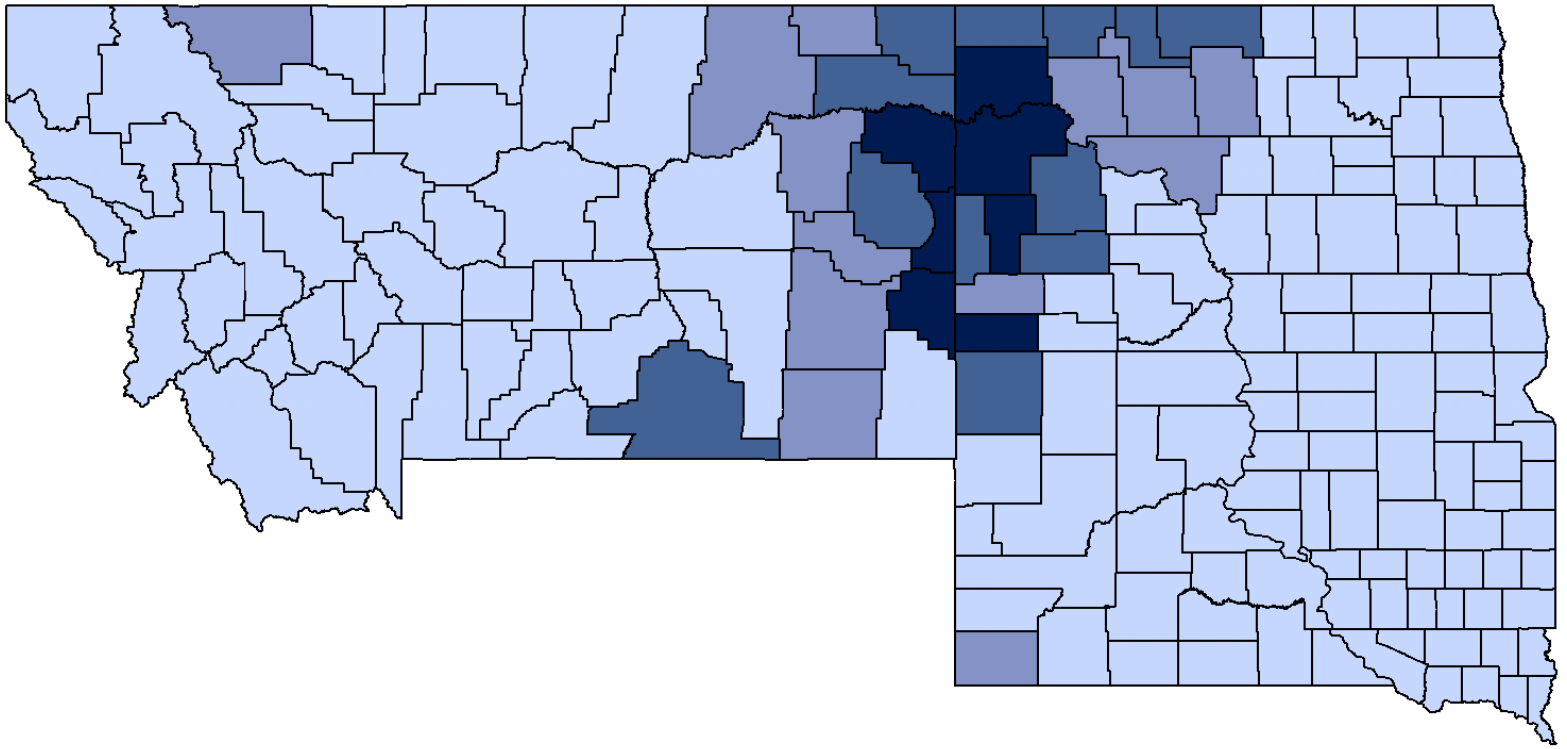
Two Econometric Challenges

- A second challenge is that at the local labor-market level, earnings, employment, and migration are **jointly determined**
- Such endogeneity could **bias** my estimates of β
- To circumvent this, I use an **instrumental variable (IV)** estimation strategy that isolates the shocks to labor demand from factors that also directly affect labor supply and migration

IV Strategy

- Use county-by-time variation in the value of oil reserves to estimate the impact of an oil price-generated **increase in county earnings on net migration**
- To do so, I exploit **three features** of oil production in the Bakken, each of which generates an important source of identifying variation:

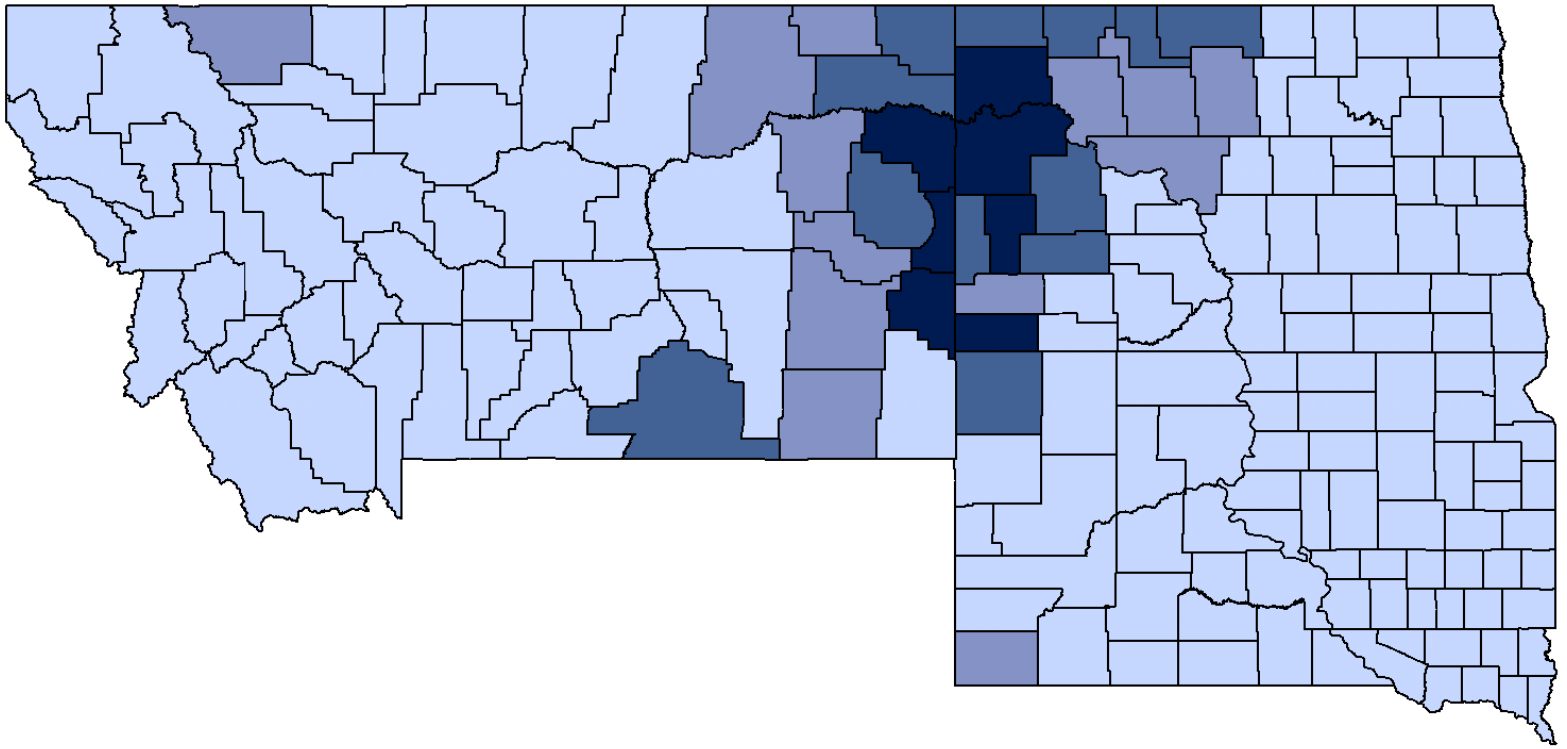
First Source of Variation: Counties Differ in their Endowments of Oil



Data – Energy Information Administration (EIA)

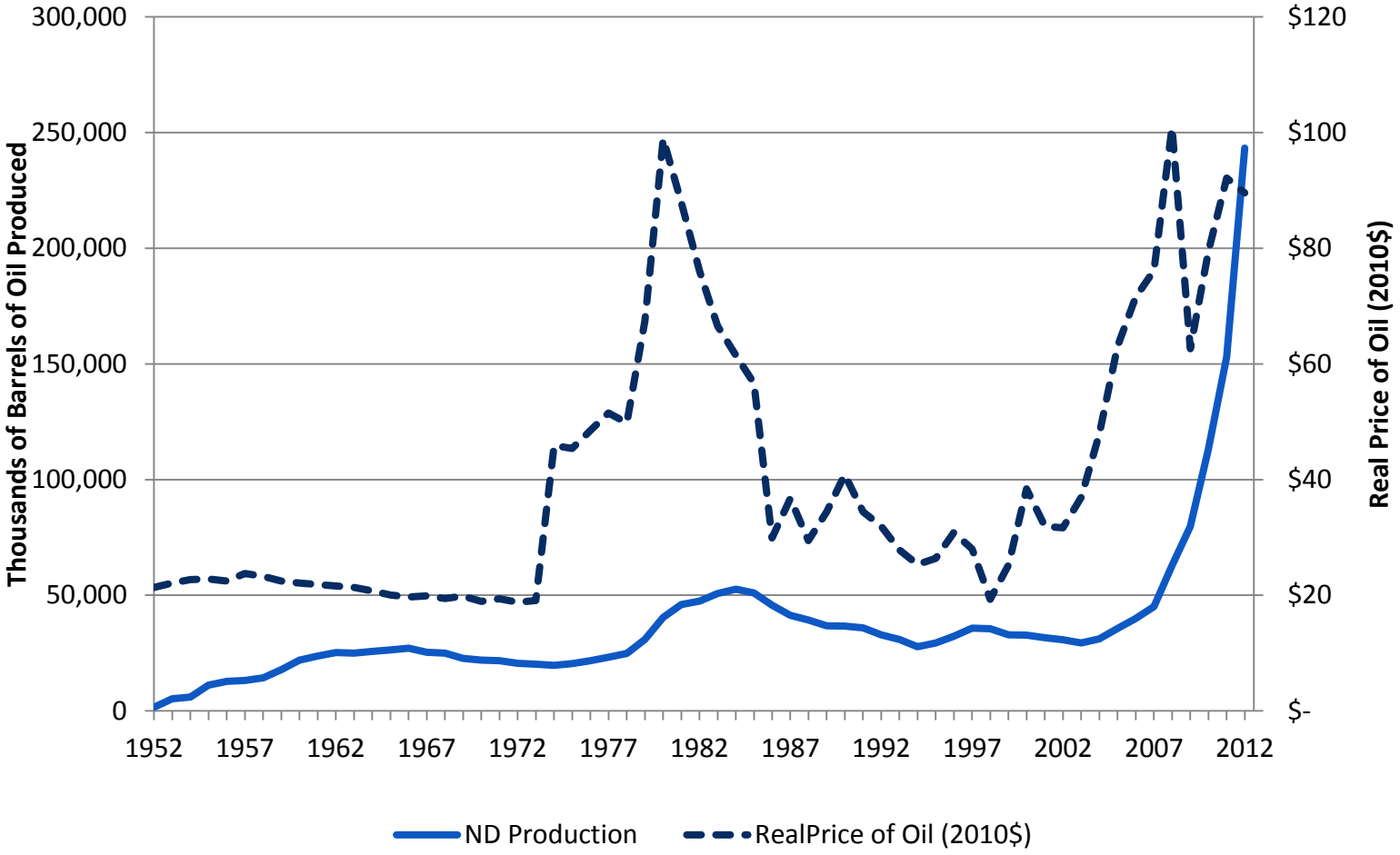
- Oil reserves from the 2004 assessment of the Bakken
- This represents reserves at the beginning of the oil boom
- Shape files of oil field reserve estimates
- Aggregate midpoint field estimates within county boundaries using MapInfo

First Source of Variation: Counties Differ in their Endowments of Oil



Second Source of Variation: Time Series Variation in Oil Prices

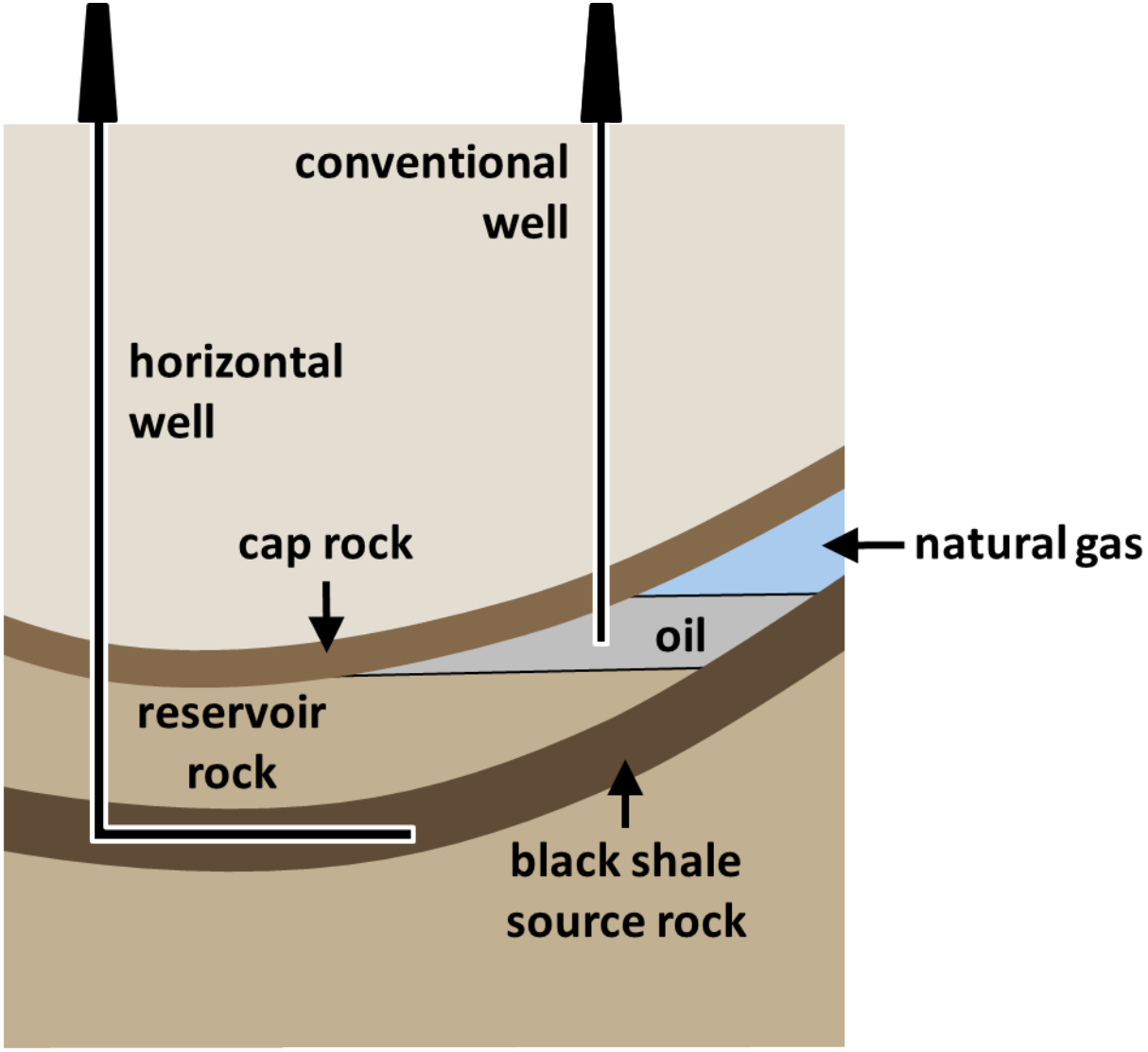
North Dakota Oil Production and the Real Price of Oil: 1952-2012



Data – Energy Information Administration (EIA)

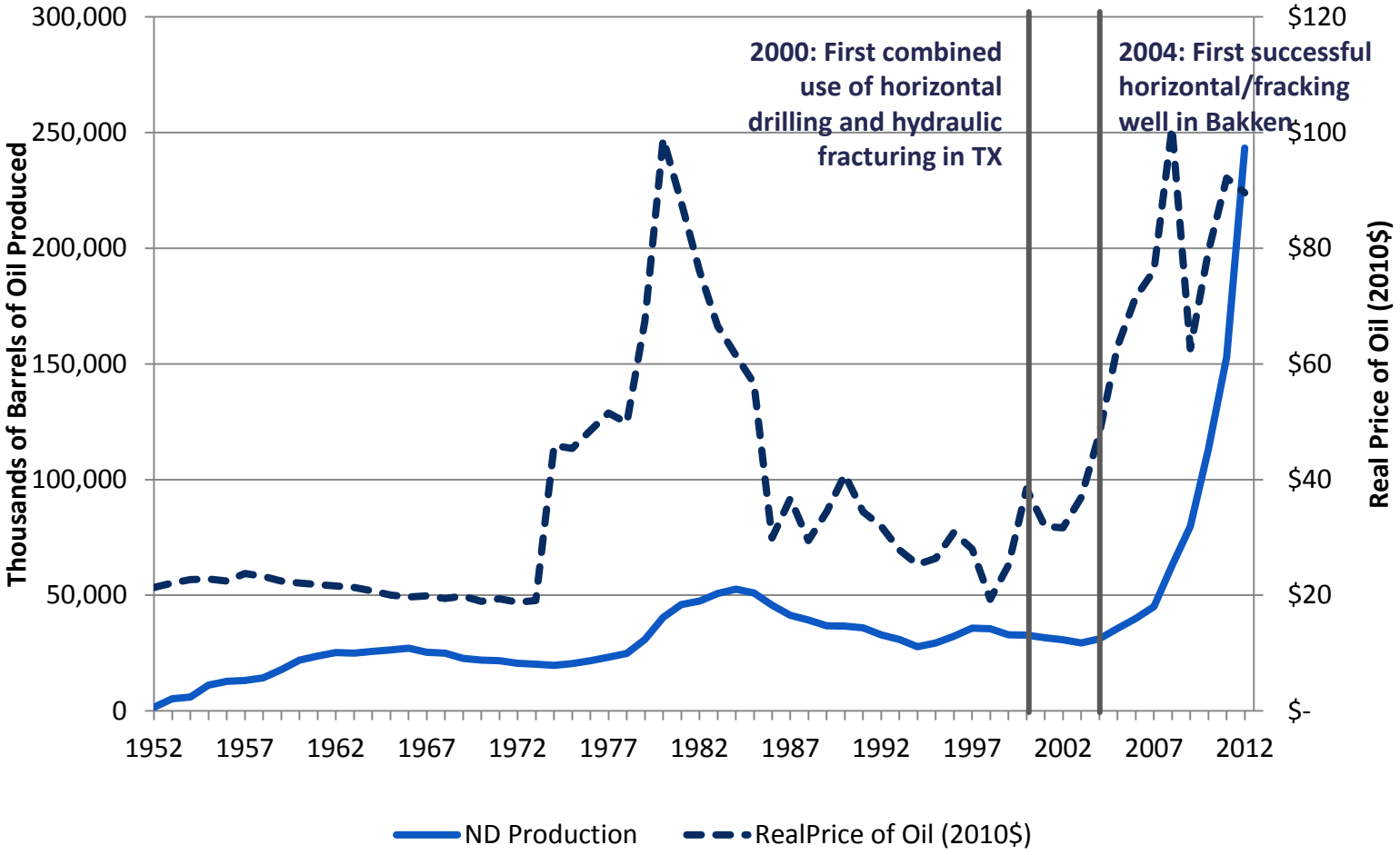
- West Texas Intermediate (WTI) crude oil prices
- Together, price data and county-level reserves generate the **county-level value of oil reserves**
- Represents county-by-time variation

Third Source of Variation: Introduction of New Extraction Technology



Third Source of Variation: Introduction of New Extraction Technology

North Dakota Oil Production and the Real Price of Oil: 1952-2012



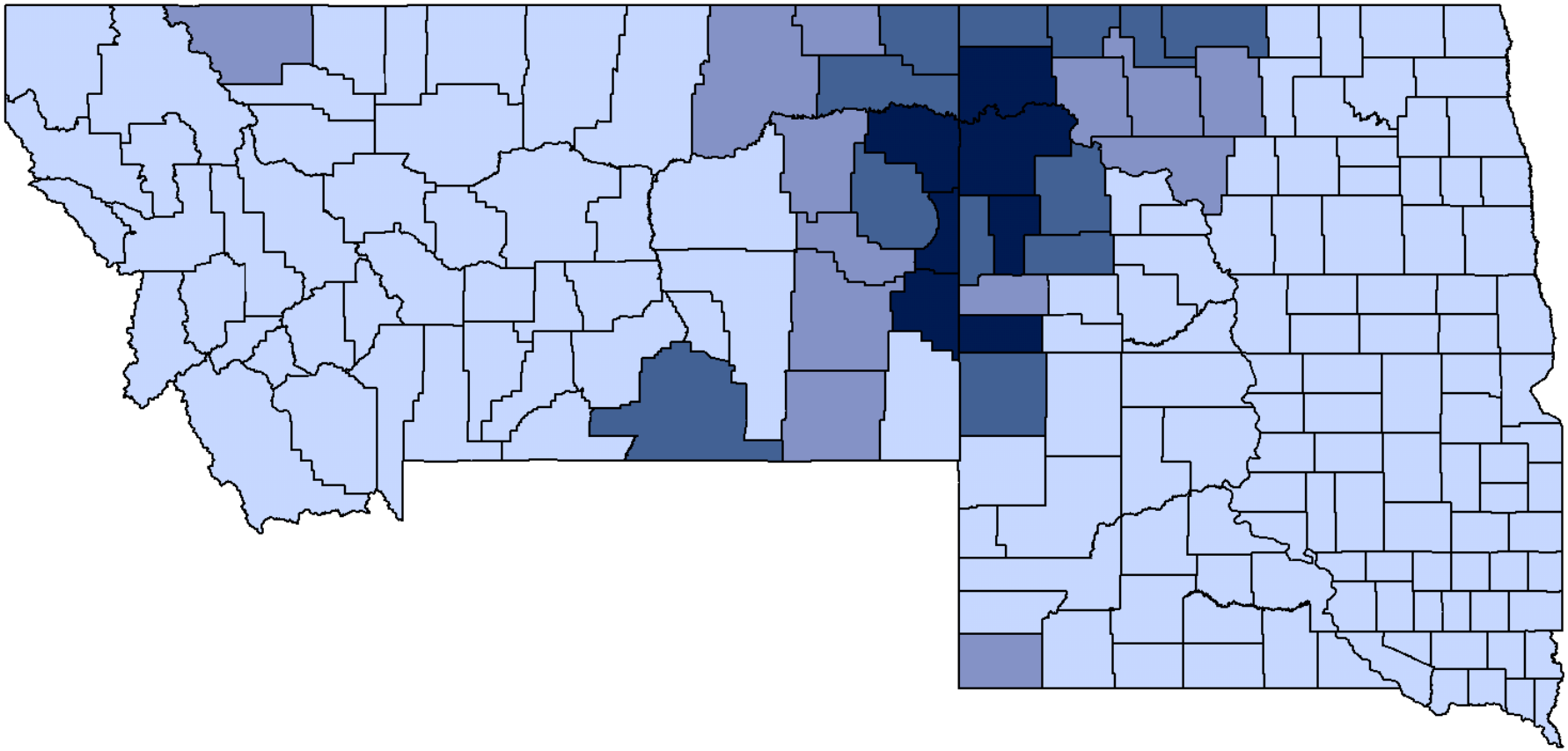
Third Source of Variation: Introduction of New Extraction Technology

- I define the boom as being associated with the introduction of these new extraction technologies

Summary and Graphical Evidence

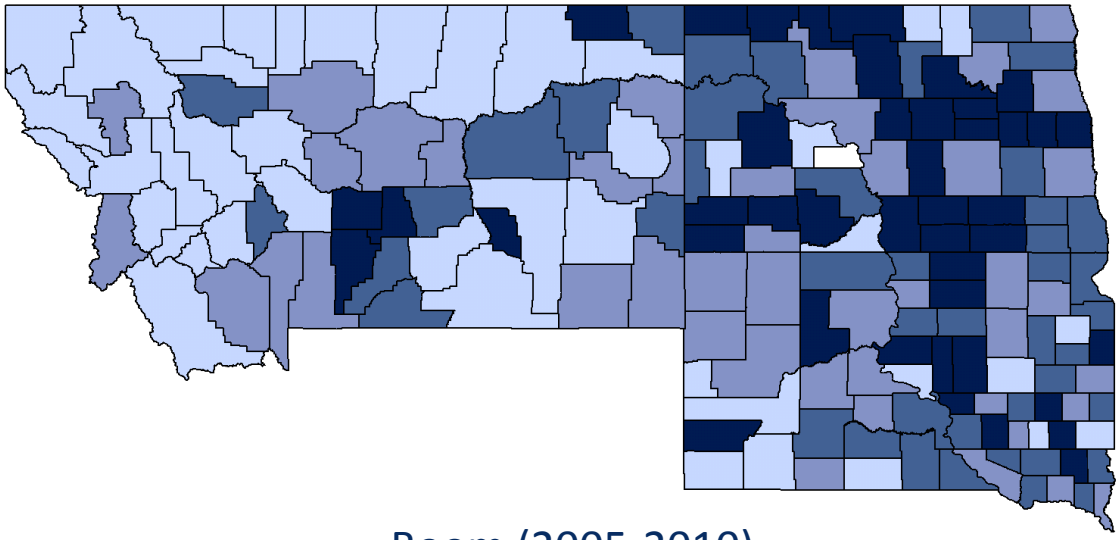
- Overall, three sources of variation that represent an interaction between:
  - Oil reserves: Variation **across space**
  - Oil prices: Variation **over time**
  - Technology: Horizontal drilling and fracking
- These generate changes in the demand for labor that differ across counties and time
- In the maps I present next, I illustrate the relationship between **oil reserves and quartiles of earnings growth** in the pre-boom and boom periods

Oil Reserves

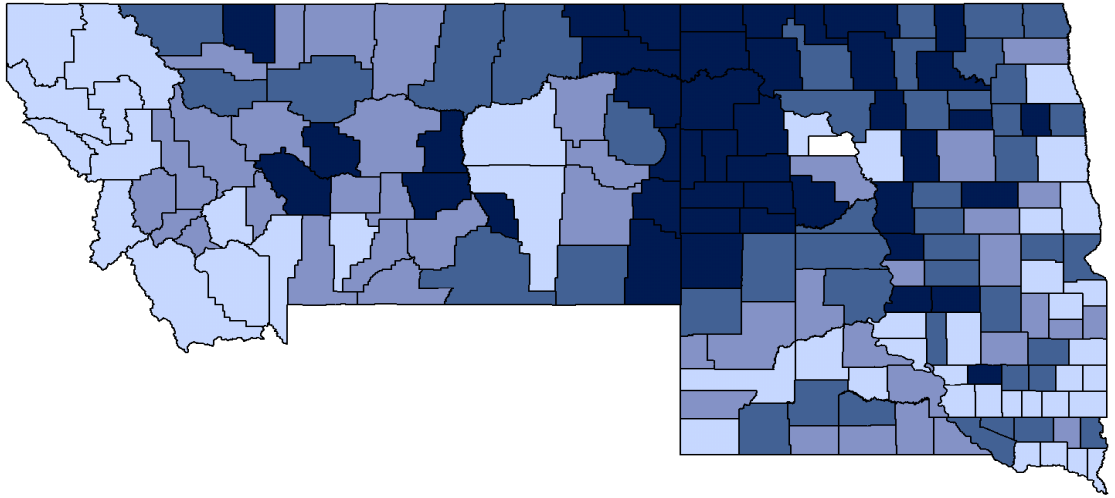


Quartiles of Earnings Growth

Pre-Boom (1993-2004)



Boom (2005-2010)



IV Strategy

- To quantify the effects shown in the maps, I use the value of oil reserves and its interaction with technology dummy as instruments for earnings

- First-stage of the IV estimation:

(3) $\Delta ln(y_{ist}) = \alpha_0 + \alpha_1 \Delta ln(v_{ist}) + \alpha_2 D^{Post\ 2004} * \Delta ln(v_{ist}) + \gamma_{st} + u_{ist}$

  - v is the value of oil reserves
  - $D^{Post\ 2004}$ is a post-2004 technology shock dummy variable

Table 2 - First-Stage Relationship between Oil Reserve Instruments and Earnings Growth: North Dakota, South Dakota, and Montana, 1993-2010

	North Dakota (1)	Three-State Region (2)
Change in the Value of Oil Reserves (α_1)	0.025	0.004
	(0.011)	(0.007)
Dummy Variable for Post-2004 x Change in the Value of Oil Reserves (α_2)	0.041	0.032
	(0.025)	(0.013)
F-Statistic	12.6	3.8
Observations	884	2,669

Notes: Robust standard errors in parentheses are clustered at the county level. All models include state-by-year fixed effects.

IV Strategy

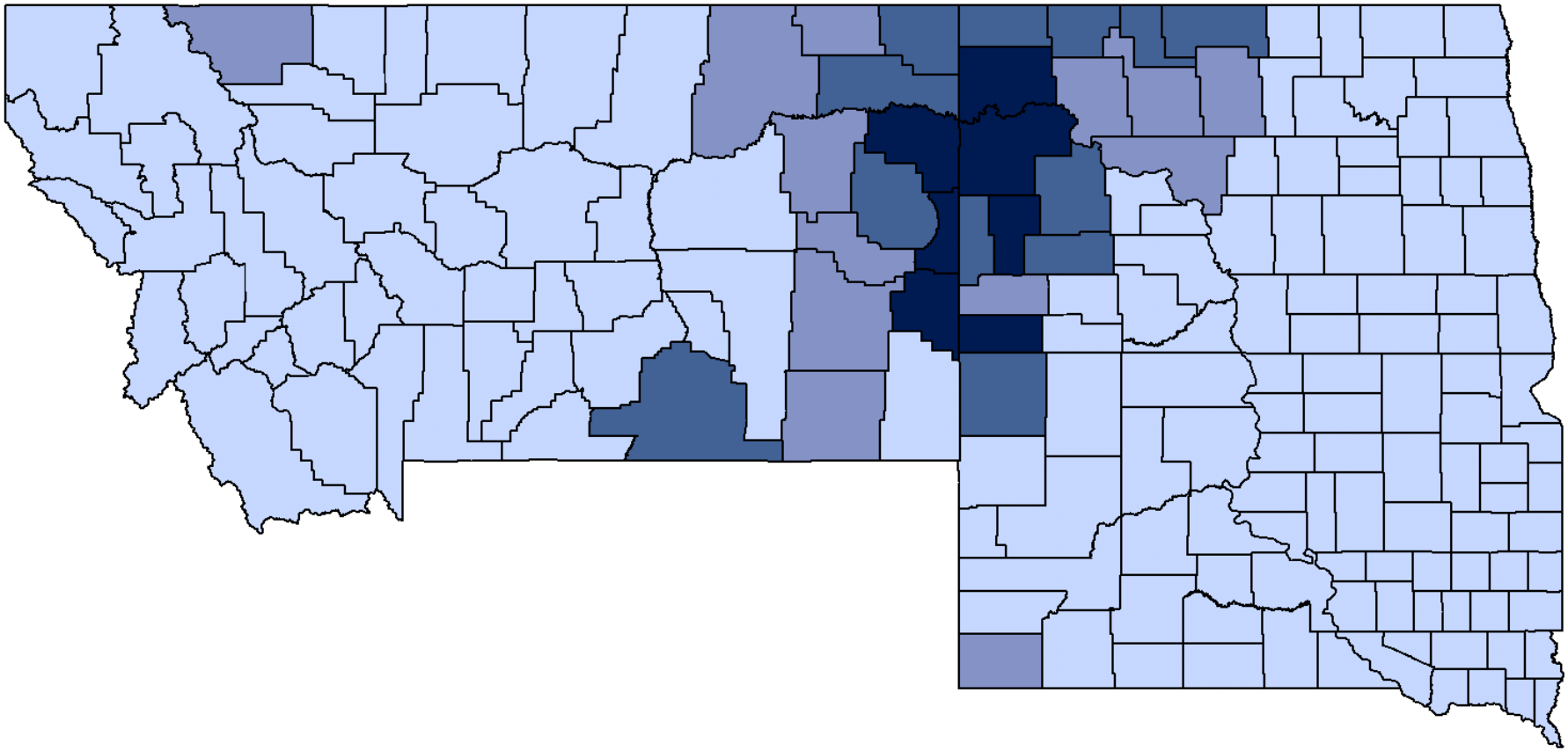
- Second stage:

(4) $\Delta m_{ist} = \delta + \beta \Delta ln(\hat{y}_{ist}) + \gamma_{st} + \sigma_{ist}$

- $\hat{\beta}_{IV}$ is the semi-elasticity of net migration with respect to earnings

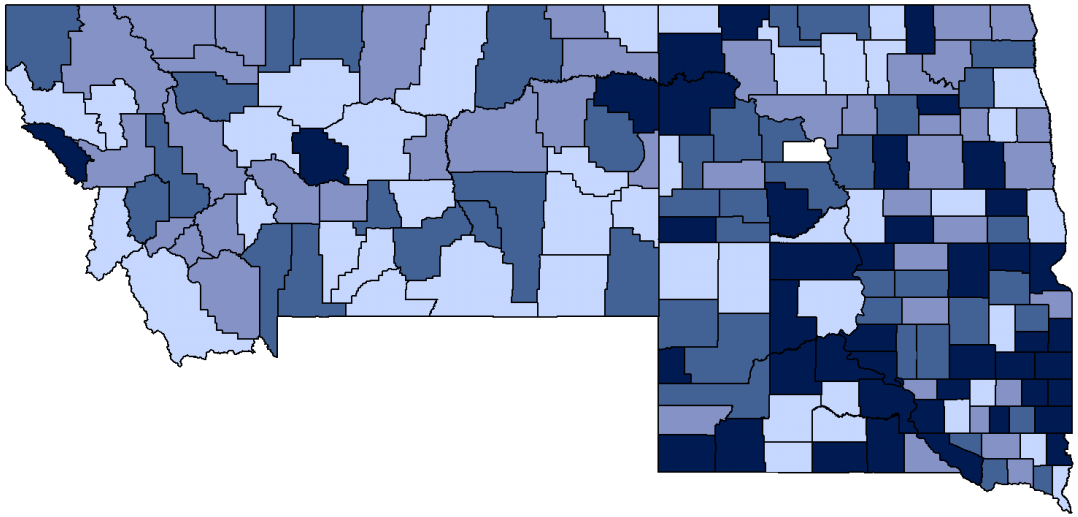
- Graphical evidence of reduced-form relationship…

Oil Reserves



Quartiles of Change in Net Migration

Pre-Boom (1993-2004)



Boom (2005-2010)

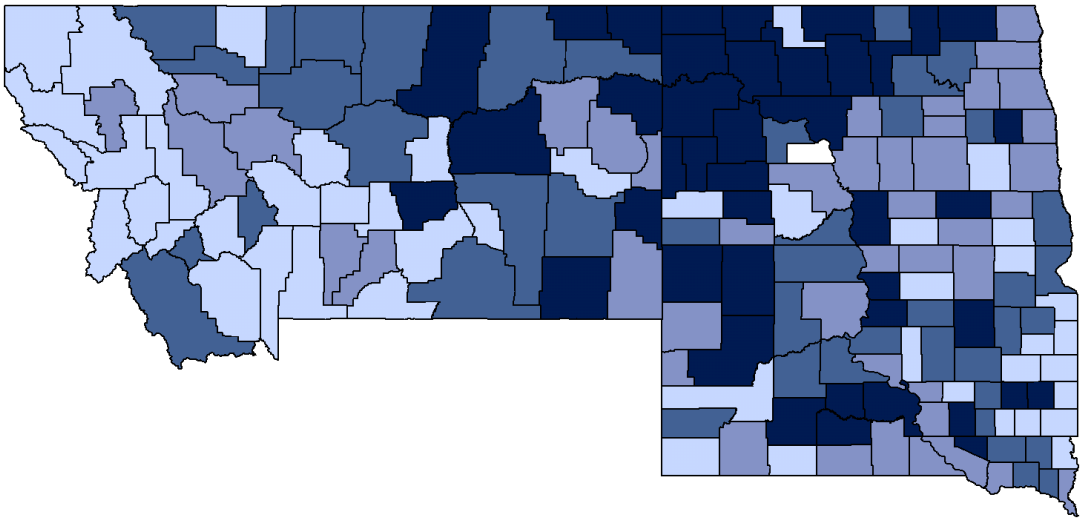


Table 3 - IV Estimates of the Impact of Earnings Growth on Net Migration: Montana, North Dakota, and South Dakota, 1993-2010

	North Dakota (1)	Three-State Region (2)
Earnings Growth	0.214	0.429
	(0.115)	(0.222)
Observations	884	2,669

Notes: Robust standard errors in parentheses are clustered at the county level. All models include state-by-year fixed effects.

- Average earnings growth
  - North Dakota: 13%
  - Three-state Region: 8%

Temporary Migration

- The IRS data used in this paper measure those who filed federal income tax returns as residents of the three-state region
- Permanent migrants should be **less elastic** in their response to changes in earnings than temporary migrants, as the fixed costs associated with a permanent move are relatively high
- I find that permanent migrants are responsive to labor income of permanent residents

Temporary Migration

- How responsive is migration to various sources of income?
  1. Consumption represents expectations of permanent income for permanent as well as temporary workers
  2. Non-labor income is likely to be a reflection of aggregate economic activity, including temporary workers
- Wedge between estimates using labor v. permanent/non-labor income provides **suggestive evidence of temporary migrants**

Table 4 - IV Estimates of the Impact of Various Measures of Economic Activity on Net Migration: North Dakota, 1999-2010

	(1)	(2)	(3)	(4)
Consumption	0.053 (0.026)			
Earnings		0.218 (0.117)		
Non-Labor Income			0.042 (0.021)	
Adjusted Gross Income				0.066 (0.035)
Observations	628	681	681	681

- Wedge between (1) and (2) provides suggestive evidence of **temporary migration**
- Estimates in (2) through (4) suggest that permanent migrants **responsive to changes in labor income**

Costs of Migration

- Previous models of migration assume that **moving is costly**
- To my knowledge, existing literature provides no estimates of migration costs

Costs of Migration

- From the basic theory,

$$\beta = \eta_A^S - \eta_A^D - \eta_B^S \left(\frac{S_B}{S_A}\right)\left(1 + \frac{C}{Y_B}\right)$$

- Define **θ as the earnings premium, $\frac{C}{Y_B}$**, paid to workers to compensate them for the cost of migrating
- Solve for θ:

$$\theta = \frac{-\beta - \eta_A^D + \eta_A^S}{\eta_B^S \left(\frac{S_B}{S_A}\right)} - 1$$

Costs of Migration

- θ is a function of:
  - The migration semi-elasticity, β
  - The labor supply and demand elasticities in each region
  - The population ratio
- To calibrate this, I use the estimate of β and make reasonable assumptions about the labor supply and demand elasticities and which areas comprise Regions A and B

Costs of Migration

- Uncompensated elasticity of labor supply: η^S=0.1
- Industry-weighted elasticity of labor demand:
  - η^D = -0.92 = (-1.3)(0.24) + (-0.8)(0.76)
- Population ratio=5
  - Region A: Western ND
  - Region B: MT, SD

Costs of Migration

- θ=0.64
- Workers require a **64 percent earnings premium** if they are to migrate to western North Dakota

Summary

- Earnings growth in oil counties significantly **increases** net migration
- **Semi-elasticity** of net migration with respect to earnings is 0.2 for North Dakota
  - 13 percent increase in earnings led to 2.6 percentage point increase in net migration rate
- Earnings premium to compensate workers for migrating to North Dakota is **64 percent**

Contributions

- Use of a natural experiment is a novel approach that provides **new causal estimates** of this relationship
- **Contributes to growing recent literature** examining the relationship between natural resources and various labor market outcomes
- Estimates of **migration costs**

Recent Decline in Prices

- Prices have **fallen by more than 50%** since June, 2014
  - Current WTI price is approximately $42 per barrel
- Recent decline represents an adverse labor demand shock that will **reduce employment and earnings**
- Back-of-the-envelope calculations suggest that this decrease in prices will:
  - Reduce earnings by 3.25%
  - Reduce net migration by 0.65 percentage points

Future Research and Extensions

- Expand sample beyond three-state region
  - Potential difficulties in regions with oil and natural gas reserves
- Temporary migration
  - Census Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES)

Thank you.

mcvachon@lsu.edu

Oil Reserves: MT, ND, SD