Empirical Macro And Finance

Atif Mian

Princeton University

Empirical Macro-Finance: Big Picture

☐ How (and why) to write an empirical macro-finance?
- o What comes first, "data" or "idea"?
  - Idea → Data

  - Idea → Data → Idea+

- o Data remains a big bottleneck in the idea generation process.
  - Empirical Macro-Finance has great potential to push forward a new research agenda.
  - This lecture provides motivation and some examples.

- What question to ask?
  - o It roughly takes the same amount of time and effort to write a paper on a "big" question versus a "small" question. So you might as well aim big …
  - o Macro asks the biggest questions
    - Why some country grow and others do not?
    - Where does business cycle fluctuation come from, and can anything be done about it?
    - Why did the U.S. have a spectacular collapse in output and employment in 2008?
  - o There is endless chatter on these issues in talk shows and op-eds
    - Data offers an opportunity to actually get closer to the truth

Example Question: Why did the U.S. have a spectacular collapse in output and employment in 2008?
- The Idea Part:
  - Think of all possible answers we know of theoretically
  - Theories may be imperfect, and we may be completely missing the true theory, but data will help us identify such deficiencies.
- Legitimate theories of U.S. collapse
  - A technological / productivity shock (e.g. RBC models)
  - An uncertainty shock (e.g. Nick Bloom and others)
  - A bank lending channel shock (e.g. BG, Kiyotaki-Moore)
  - An aggregate demand shock (e.g. Eggertsson-Krugman)
- We need to fully specify these theories before we can taken them to data.
  - Notice the first two explanations have little finance in them.
  - The latter two posit finance to be the driver of macro changes
  - Does finance matter for macro?

Macro-Finance Nexus

☐ Does finance matter for macro? (I define finance as the act of
  borrowing/lending)

  o Not necessarily. Many macro models happily attempt to explain
    the world without any reference to finance

    - For example, take the basic representative agent model.
    - Representative agent cannot borrow or lend to himself. Hence
      no role for finance.
    - Efficient financial markets, both in terms of financing needs
      and risk-sharing implicitly embedded in such models.
    - The focus in these models is on non-financial factors, such as
      technological shocks, population growth, uncertainty,

information aggregation, inflation expectations, price and wage rigidity.
- Finance is largely irrelevant in these models, it is essentially just a side-show.

Why should finance matter for macro?

    o We have to get out of the representative agent framework

        - Heterogeneity matters

        - Distribution matters

    o Which heterogeneity to focus on?

Entrepreneurs versus savers: The investment channel
- o Only a few Bill Gates and Jeff Bezos with high marginal product of capital.
- o Want to match entrepreneurs with savers (who have low MPK), but:
    - People lie about project MPK (screening problem)
    - People lie about project cash flows (monitoring problem)
- o Hence need an intermediary to screen, monitor and enforce. But who monitors the monitor?
    - Therefore, intermediary must put up capital of its own,
    - => Investment = f(Intermediary Capital)
- o Now finance matters for macro via the investment / bank lending channel.

High versus low MPC: The consumption / aggregate demand channel.

- Life cycle hypothesis says consumers differ over time in their marginal propensity to consume => need to borrow and lend to each other.
- Some agents may be more impatient than others => need to borrow.
- Borrowing households need net-worth to be able to borrow and consume,
  - => Consumption = f(Borrower's Net Worth)
  - The channel from borrowers' consumption response to aggregate output and employment is more nuanced (e.g. Eggertsson-Krugman)

Summary: Macro-Finance Nexus
- Financial shocks (e.g. asset price movements) in combination with leverage shift the distribution of net-worth across agents
- The redistribution of net-worth impacts the real economy either via the "investment channel", or via the "consumption channel".
- Note that the two channels are very different in terms of their policy prescription: we may end up recapitalizing the wrong sector.
- Investment Channel

Firms ← Banks ← Households

- Consumption Channel

Firms ← Banks → Borrowing Households

Banks ← Lending Households
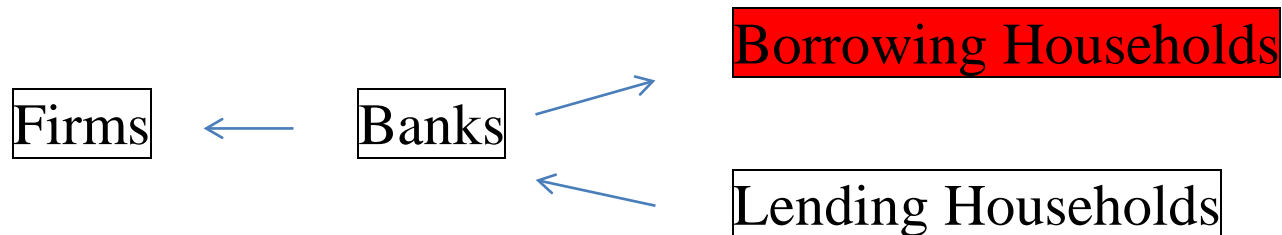
Macro-Finance Empirics

☐ Example Question: Why did the U.S. have a spectacular collapse in output and employment in 2008?
- o We want to test the empirical validity of each theory
- o We want to quantify the importance of empirically relevant hypotheses.

Start with deriving testable predictions of each theory, and focus on the
central / robust implications of each theory
   o For example, the investment channel predicts
      - Firm investment declines more for bank-dependent firms
      - Firm investment declines more firms reliant on harder-hit
        banks
   o For example, the consumption channel predicts
      - Household experiencing a larger reduction in net-worth cut
        back more on consumption
      - Employment linked with the above consumption declines
        more.
   o We are now beginning to give sufficient shape to our ideas to go to
     data
      - Next step, empirical challenges.

Methodological Issues

  o Identification (relatedly, think carefully about statistical power)
  o General Equilibrium effects
      - "local"
      - "global"
  o Quantification (relatedly, think carefully about appropriate
    structural model)
  o Simplicity and Transparency
      - Worry about empirical model misspecification issues (more
        "non-parametric" the better)
      - Take the "con" out of econometrics (Leamer)

☐ Identification
  o RCT might be the benchmark, but it is not a realistic target in the macro-finance space
  o Look for natural experiments, instruments, regression discontinuity …
  o Most important: fully specify your exclusion restriction, and "test" it
    - No empirical paper is ever written without an important set of assumptions under which its results are valid.
    - You must specify these assumptions explicitly
    - You must test these assumptions on all possible observables (you can never completely test the exclusion restrictions, but you can always do something)
  o Keep in mind that identification is related to the notion of statistical power, i.e. your ability to reject a hypothesis if it is indeed false.

- Thus refrain from claiming victory after results that might be "consistent" with your hypothesis but do not *uniquely* identify your hypothesis.
- This is a common problem in the empirical macro space.
- If the best you can do is show that the data is "consistent" with your hypothesis, then at the minimum run through all possible alternative hypotheses and tell the reader explicitly the likelihood of finding the same result under each of the alternative hypotheses.

General Equilibrium Effects

- o Identification is just the start. You want to be very careful about general equilibrium effects when investigating macro questions. There are two kinds of general equilibrium effects to worry about.
- o The "local" general equilibrium adjustment done at the level of the household, bank or firm.
  - For example, if my bank faces a large reduction in capital, it might very well cut credit to me (the investment channel is validated). However, such an investment channel may have no macro impact if I can easily tap into my internal funds to keep my investment constant, or go to another bank that does not face such constraints. (see Khwaja and Mian AER 2008 for an example of estimating such general equilibrium effects)
- o The "global" general equilibrium effect due to economy wide adjustments

- For example, the borrowing household may cut back consumption in response to a negative house price shock. But this may not result in a reduction in aggregate demand if the drop in borrowers' consumption leads to lower interest rates which in turn convince the lending households to consume more. (see Mian and Sufi 2012 for an example of testing whether such an effect exists)

☐ Quantification
- o So you have your well-identified estimates. How do you quantify the macro impact?
- o Understand that most empirical estimations estimate some form of "local average treatment effect".
  - You must know fully which "local average" you are estimating
  - Understand any limitations this imposes on "extrapolating" your coefficients to the entire population.
- o Think carefully about in-sample versus out-of-sample predictions.
  - For example, say you have estimated the causal impact of house price growth on leverage. How much of total leverage growth can be explained by your coefficient? You want to use only the "in sample" house price variation when integrating up the cumulative effect of your coefficient. (see Mian and Sufi AER 2010 for an example)

o No empirical strategy using cross-sectional variation can identify the "level effect", since the level effect is a component of the overall time effect which is potentially polluted by every other hypothesis floating around.

- This is where structural models can be really useful
- You want to, (i) write down the relevant structural model, (ii) use the cross-sectional estimates to parameterize the model, and (iii) run the model with counter-factual exercises to estimate the full macro impact. (see Midrigan and Phillipon 2012 for an example)

Simplicity and Transparency
- o If you cannot show your main result in a simple graph, you may have a problem.
- o Simple and flexible / non-parametric estimation can be a lot more powerful than complicated "teched up" specifications that often rely on obscure functional form assumptions for identification.
- o Strive to make the data public, or help people gain access to it.
- o The real world is always more interesting than our theoretical prejudices
    - Don't waste time writing up "results" that do not survive rudimentary robustness checks.
    - If your data is good quality, go and revisit your theory. You will often discover something more interesting.

Data

☐ Which data to use?

  o Traditional aggregate time-series data used in macro has serious limitations to address the methodological issues raised above.

  o Value lies in going to micro data for answering macro questions. The richness and variation available in micro data provides you with a much better chance to address identification concerns, assess general equilibrium effects, and quantify the overall magnitude.

  o Why now?

    - The IT wave has exponentially increased the availability of micro data sets.

- Think about it … *every* financial and economic transaction is now recorded.
- The data we need is often sitting somewhere …. we just need to find it (and convince people to share it).
- Makes it a lot easier to think about questions and hypotheses first, and then look for data.

What is available?
- o Consumer credit bureau data
    - Information on borrowing amount, type of borrowing, default status, location, age, sex, credit limits, credit score.
    - Available at quarterly frequency since around 1991.
    - Available at individual level, or aggregated at zip code / county / MSA level.
    - Sources: Equifax, Experian, TransUnion (private)
- o House price indices
    - Available at quarterly frequency since at least 1990.
    - Available at zip code / county / MSA level.
    - Sources: FHFA (public), Case Shiller, Core Logic and Zillow (all private)
- o Mortgage application HMDA data

- Information on every mortgage application, mortgage amount, purpose, location, race, reported income, lender, approval/denial.
- Available annually from 1992 onwards.
- Source: HMDA (public)

o Mortgage data
- Mortgage amount, default status, location, other information given at time of application (LTV).
- Source: McDash / LPS (private)

o Bank and Stocks data
- Listed company financials, stock prices, bank loans, and bank financials.
- Sources: Compustat, CRSP, DealScan, and Call Reports (available at most universities via wrds)

o Employment and Payroll data
- Available at at least quarterly level since 1990.

- Broken down by industry and upto zip code level geographically.
- Source: BLS (public)

o American Community Survey
- Survey of a large random sample in the U.S., conducted every two (or three?) years
- Includes self-reported wages, hours worked, industry, demographics, education, asset ownership etc.
- Source: Census (public)

o Income Data
- Information of AGI, wages, dividends and interest income
- Annual frequency at the zip code / county / MSA level.
- Has a 3-year reporting lag
- Source: IRS (public)

o Consumption

- Quarterly data since 1996 at the zip code level on purchase of new automobiles
- Source: RL Polk (private)
- Master card Advisors: annual data on expenditures through credit card. (private)
- BuildFax: Building permits data at a fine geographical level (e.g. zip code). (private)

o Prices
- BLS provides data behind the CPI. Goods prices at the MSA level.

Results

The 2002-2006 Credit Shock

Figure 1
U.S. Household Debt

Household debt to income ratio
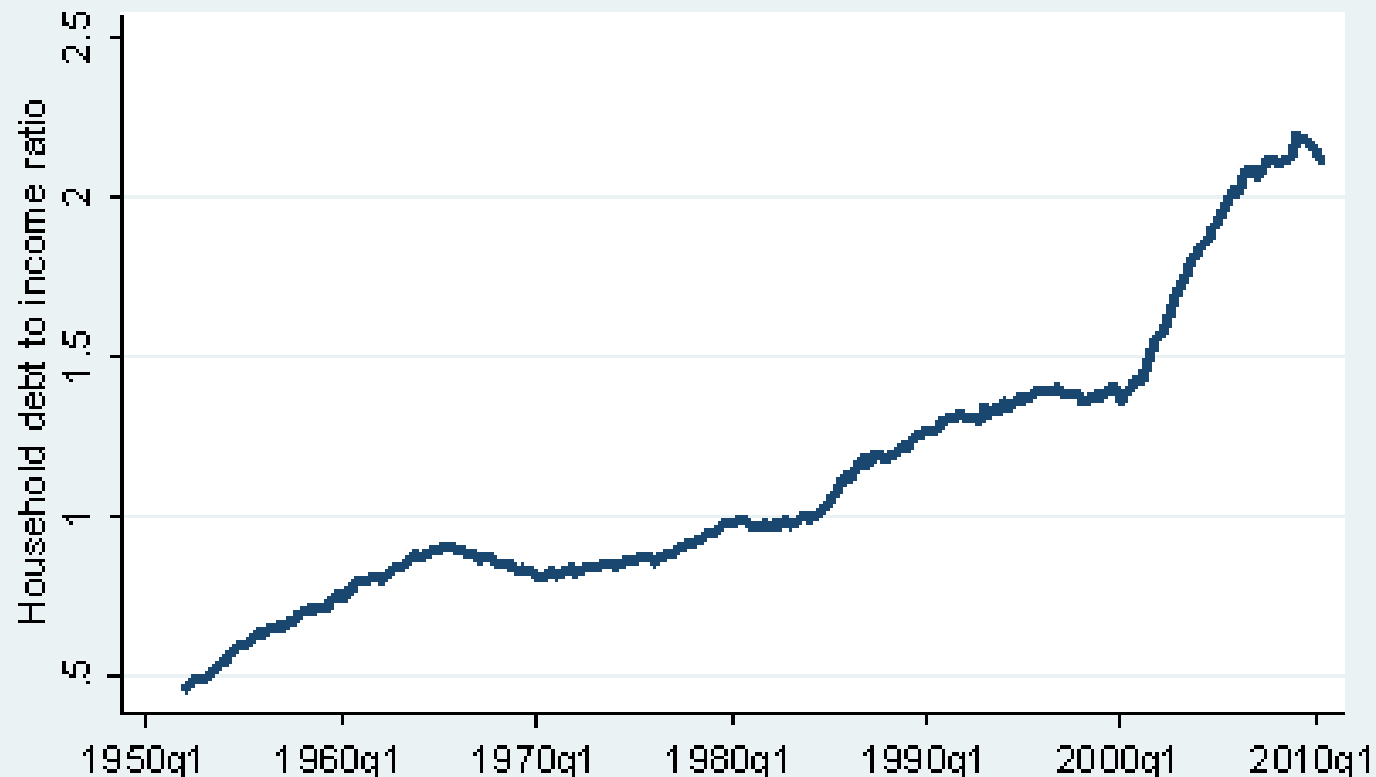
A Credit Supply Shock?
(Mian and Sufi QJE 2009)

- The market imposed a strong credit supply constraint on prospective borrowers in 1996, esp. subprime.
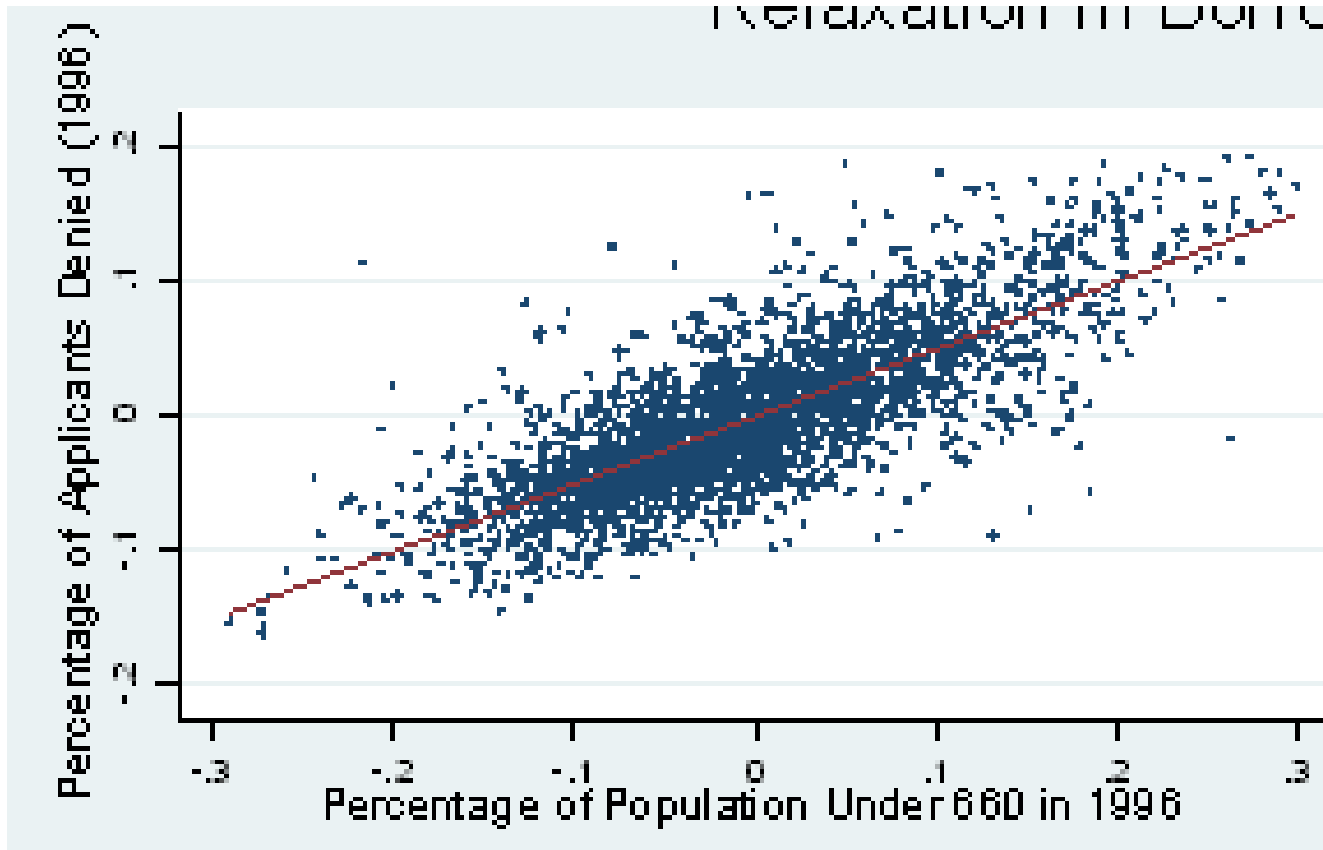
The constraint kept tightening ....
Until 2002
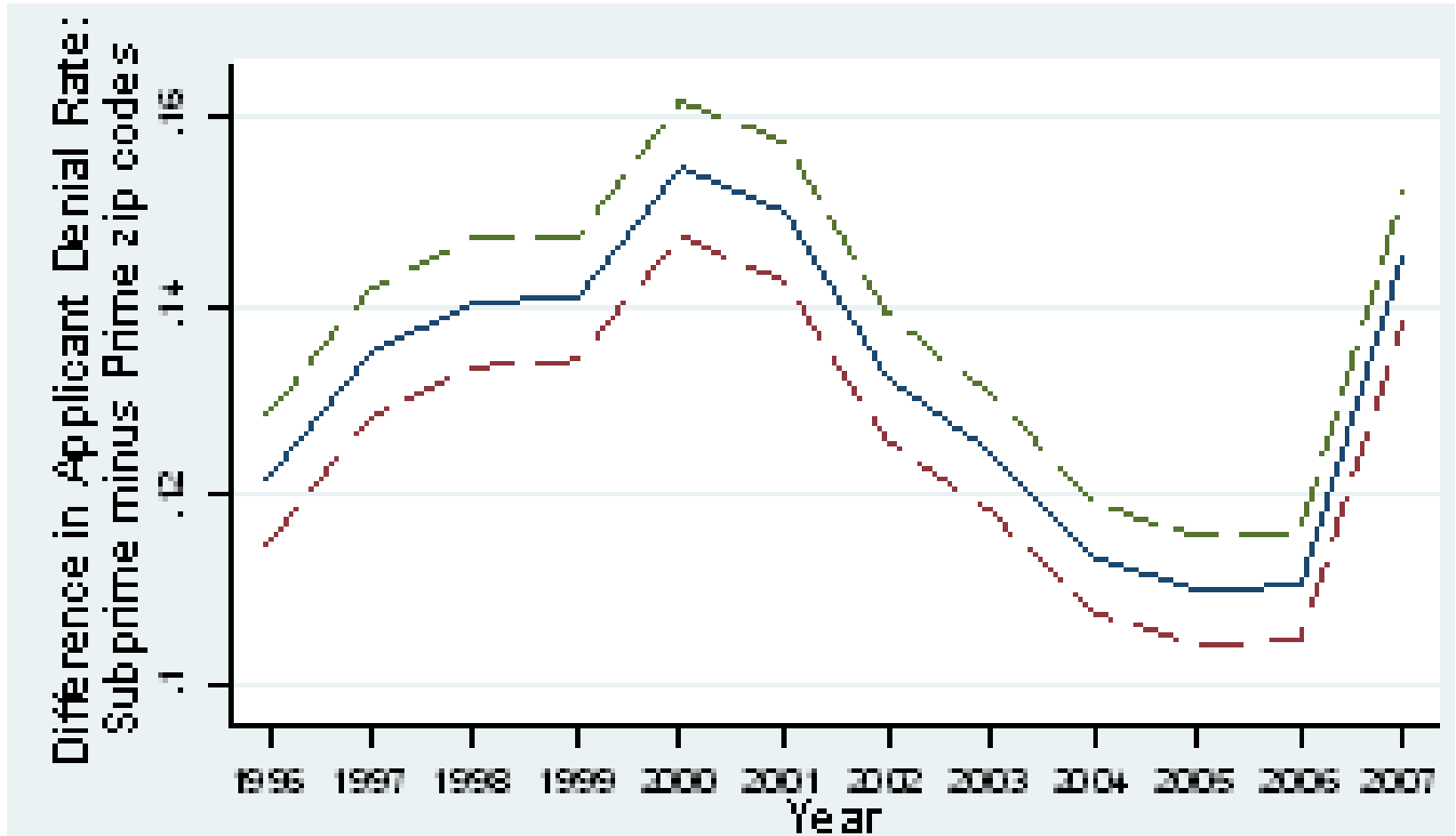
Current Account Balance
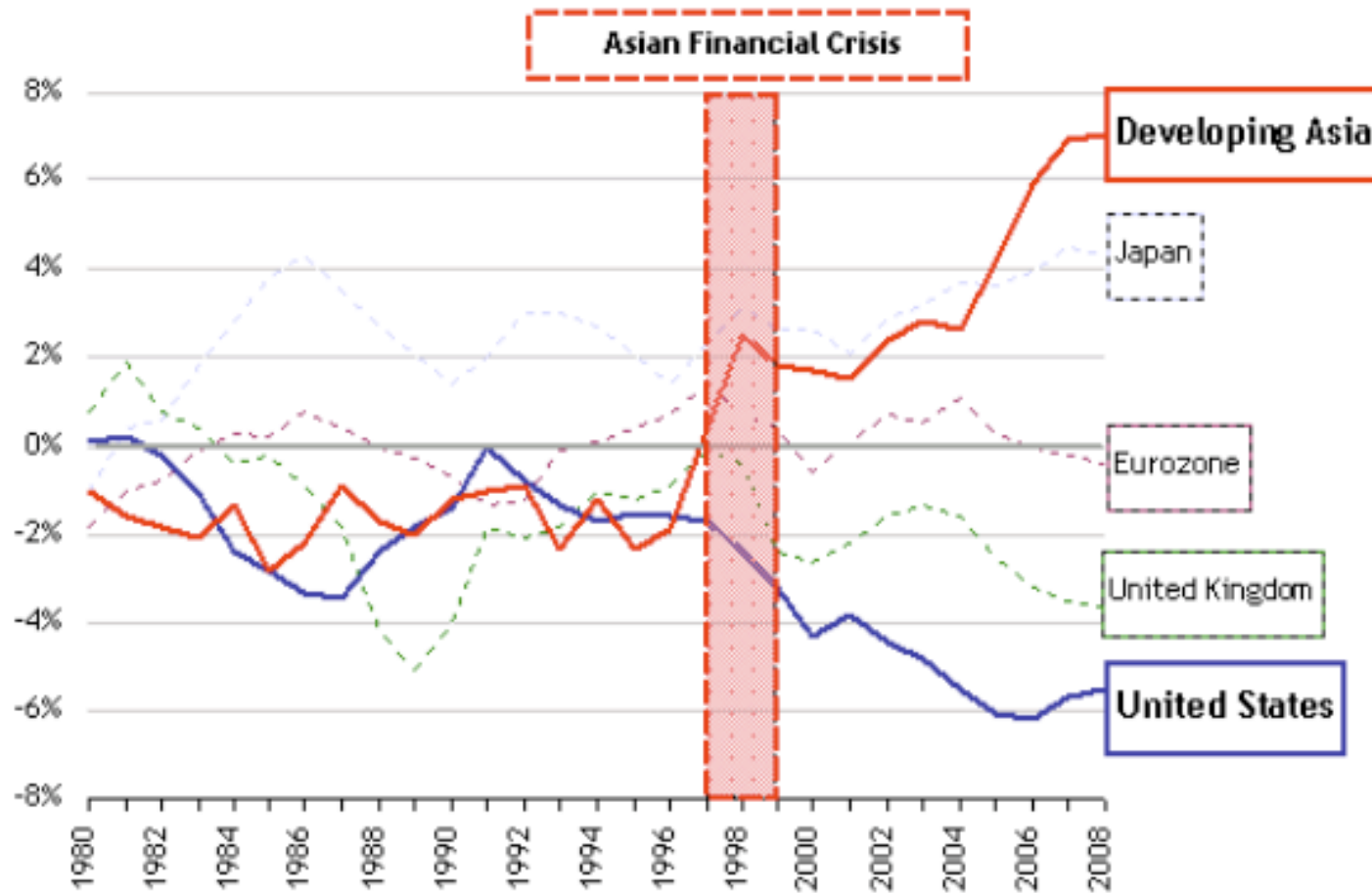
Figure 3
Credit Growth And Income Growth Over Time
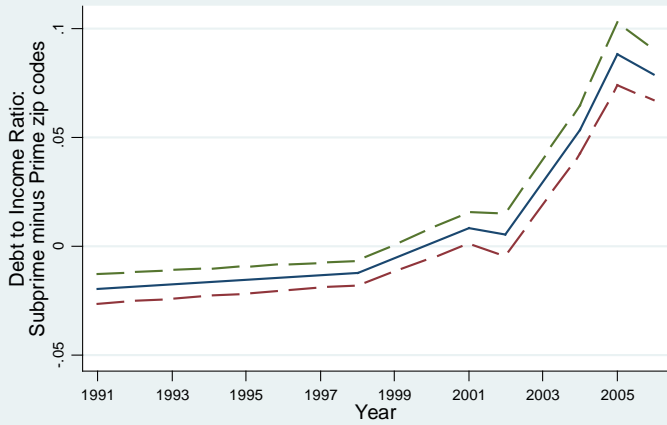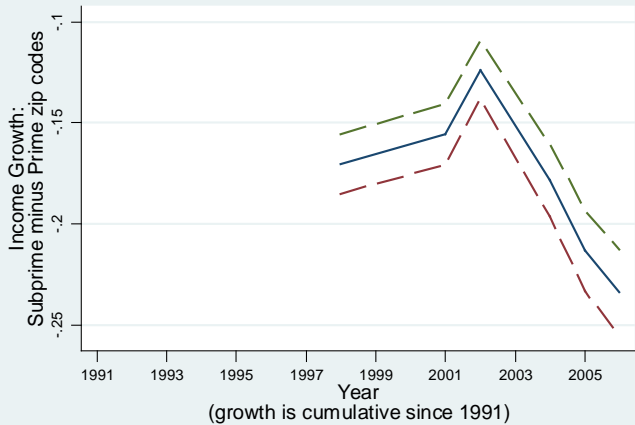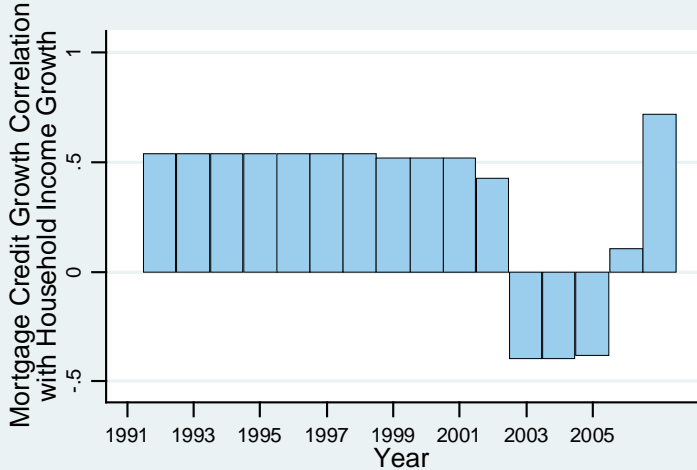
Subprime and Prime zip codes are defined to be the highest and lowest quartile zip codes within each County based on the fraction of residents with a credit score below 660 in 1991.

Instrument:
From Credit Supply to House Prices
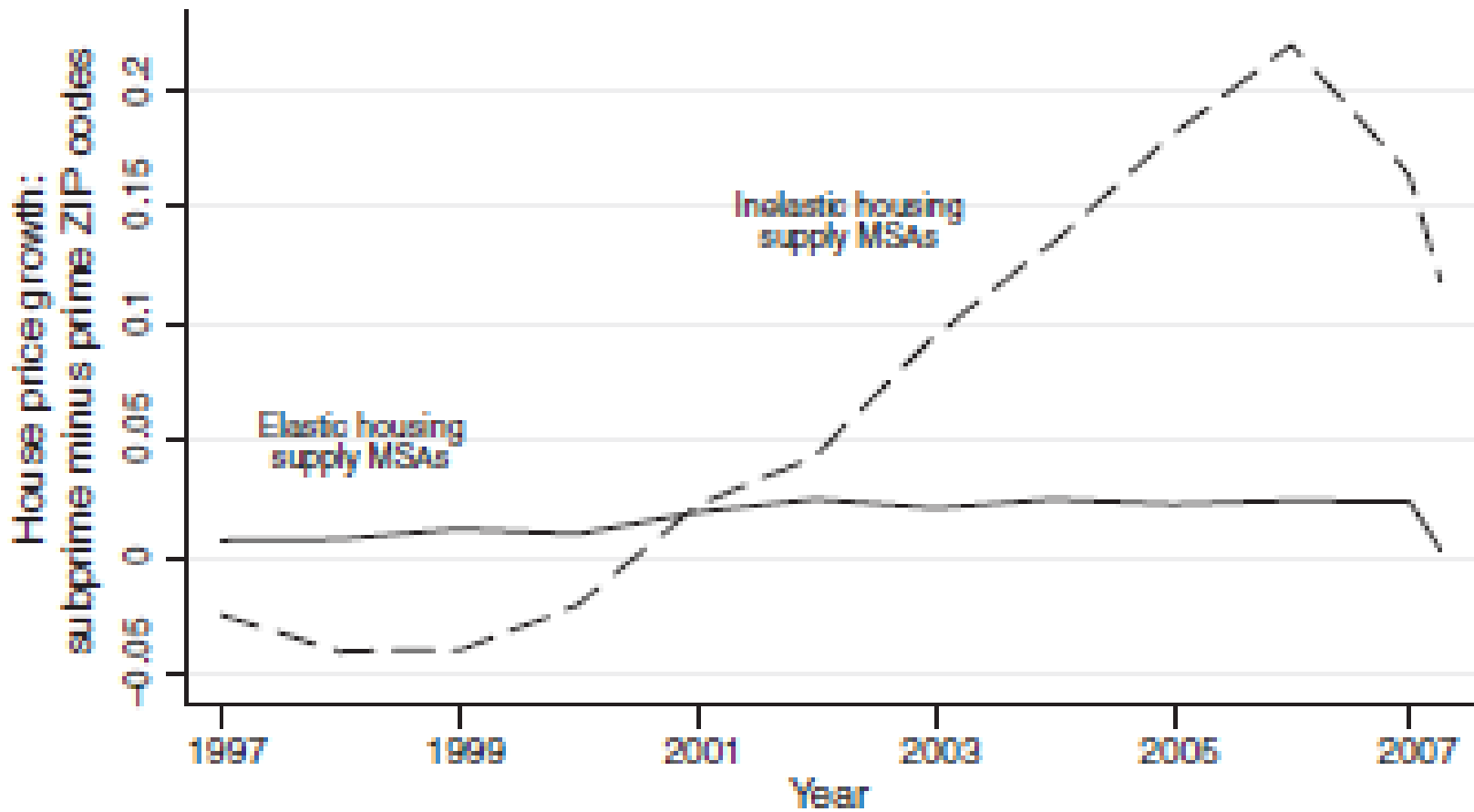
From House Prices To Leverage
(Mian and Sufi AER 2010)
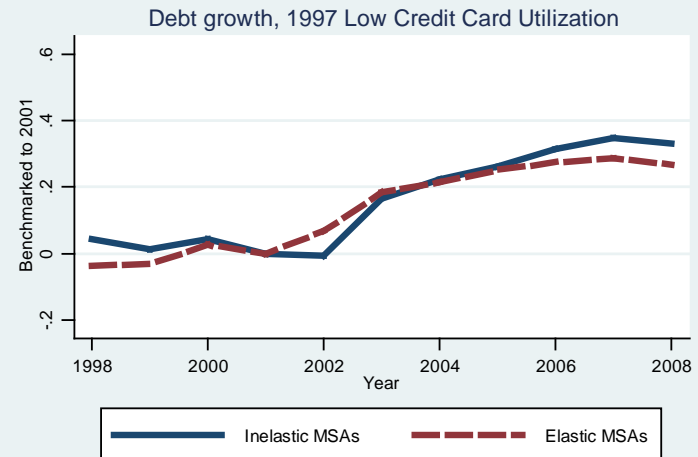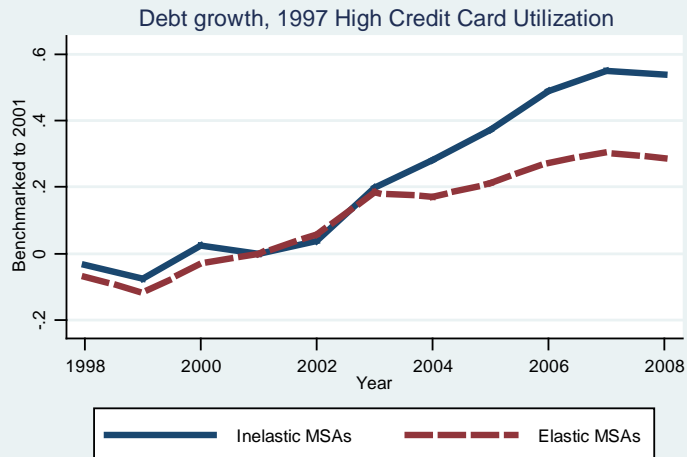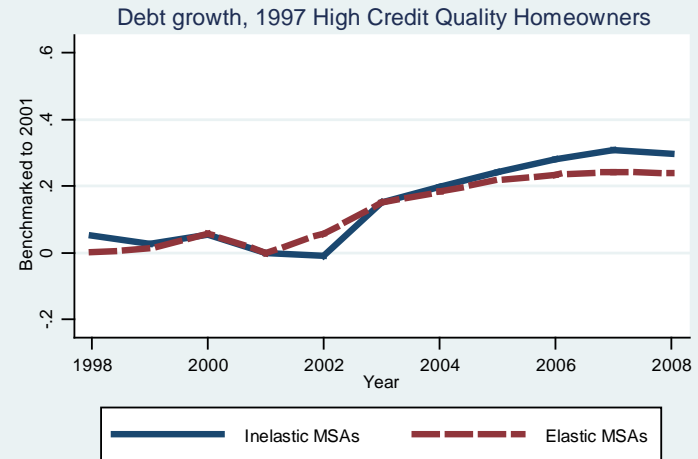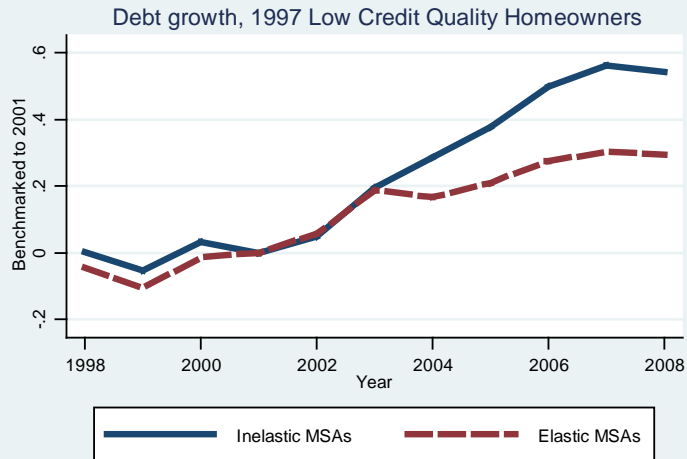
Testing Exclusion Restriction

TABLE 5—EXAMINING EXCLUSION RESTRICTION

Panel A. Income and employment measures

Left-hand side variable	Business payroll growth 2002–2006 (1)	IRS wage growth 2002–2006 (2)	Employment growth 2002–2006 (3)	Payroll growth shock (4)	Wage growth shock (5)	Employment growth shock (6)
Housing supply elasticity	−0.003 (0.006)	−0.020*** (0.003)	−0.006 (0.013)	0.016** (0.007)	−0.006 (0.005)	−0.002 (0.014)
Observations	2,241	2,306	2,241	2,220	2,306	2,220
R^2	0.00	0.04	0.00	0.00	0.00	0.00

Panel B. Home, credit card, and renters' debt

Left-hand side variable Sample	Home debt growth 2002–2006 Homeowners		Credit card balances growth 2002–2006 Homeowners		Total debt growth 2002–2006 Renters	
	(1)	(2)	(3)	(4)	(5)	(6)
Instrumented HP growth, 2002–2006	0.535*** (0.120)	2.718*** (0.617)	0.080 (0.110)	−0.139 (0.426)	−0.029 (0.161)	0.260 (1.280)
Instrumented HP growth, 2002–2006 *(Credit score, 1997)/100		−0.283*** (0.073)		0.023 (0.058)		−0.041 (0.160)
(Credit score, 1997)/100		0.093*** (0.033)		0.069*** (0.026)		−0.007 (0.081)
Observations	13,337	13,337	13,337	13,337	5,058	5,058
R^2	0.00	0.00	0.01	0.00	0.00	0.00

U.S. Household Deleveraging, Aggregate Demand and Unemployment

(Mian, Rao and Sufi 2012)
(Mian and Sufi 2012)
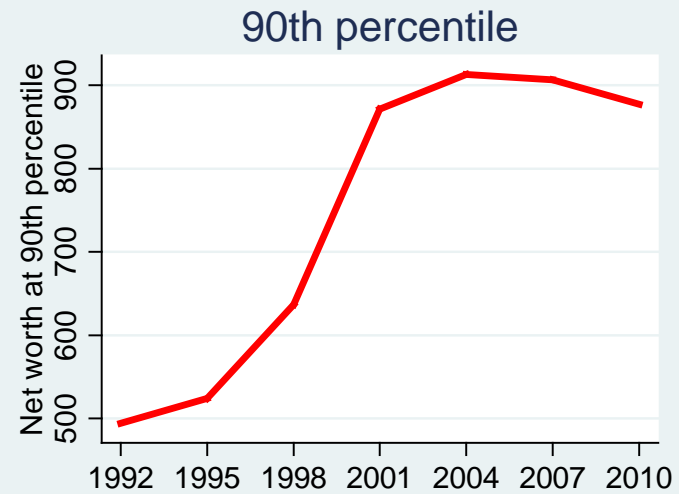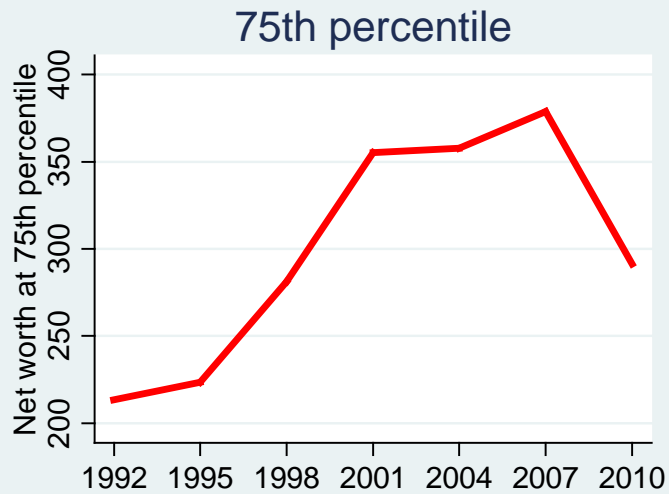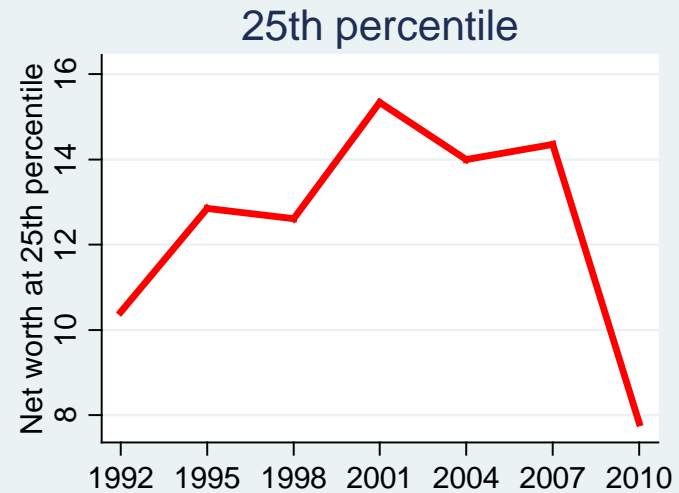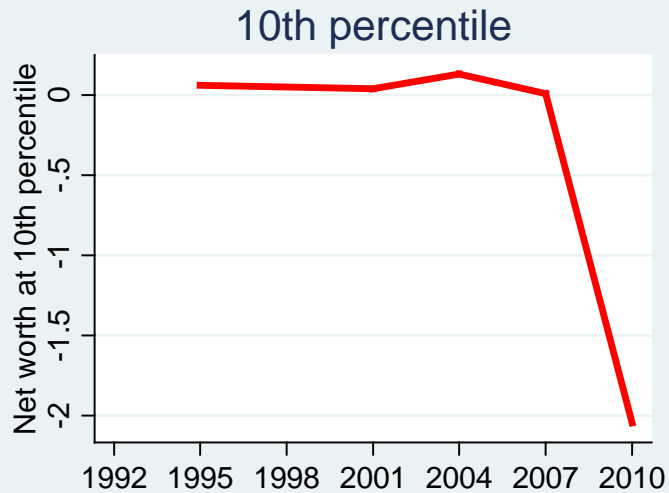
The Financial Shock

House prices
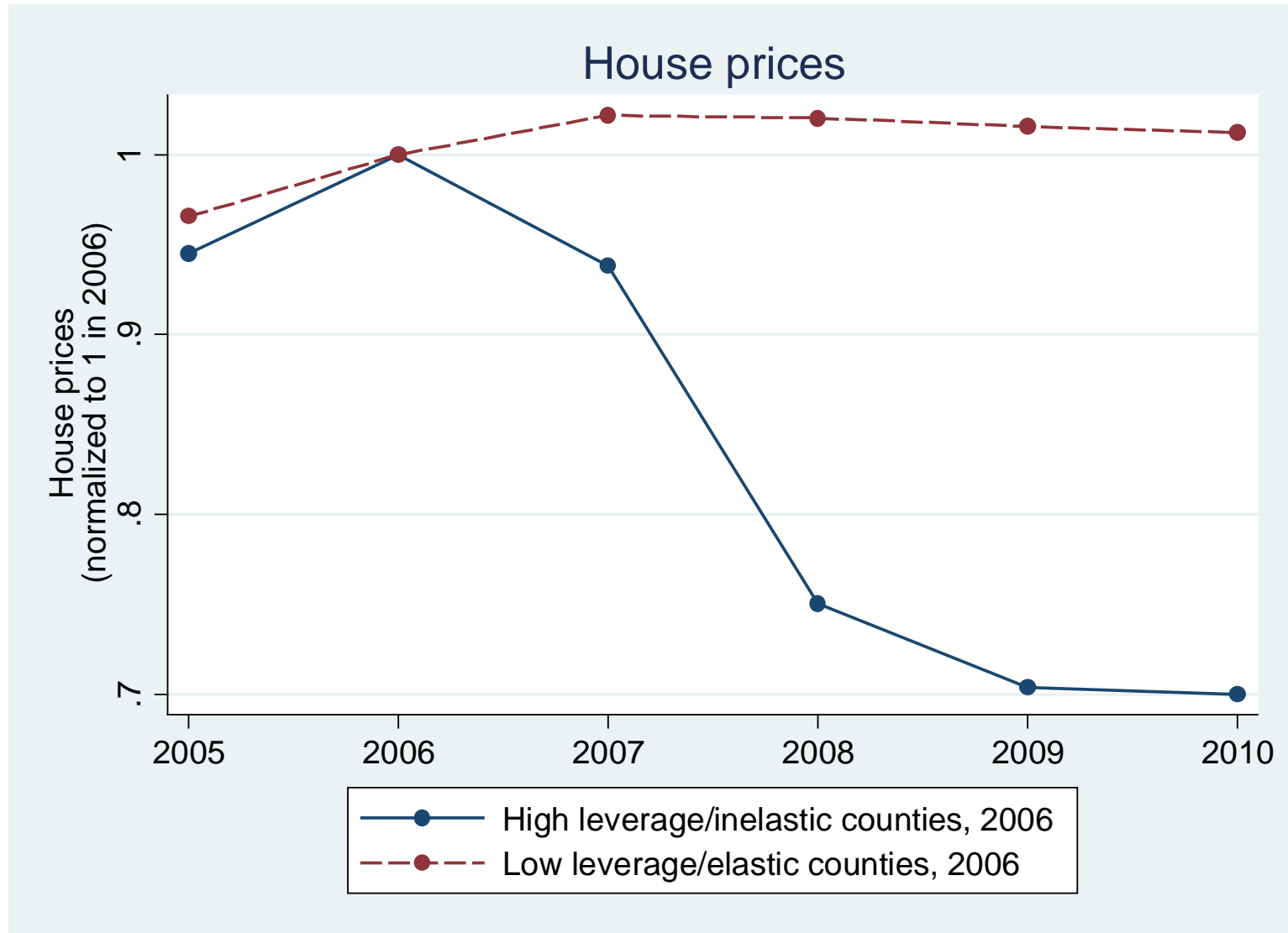
Does it translate into macro aggregates?

Does it translate into macro aggregates?
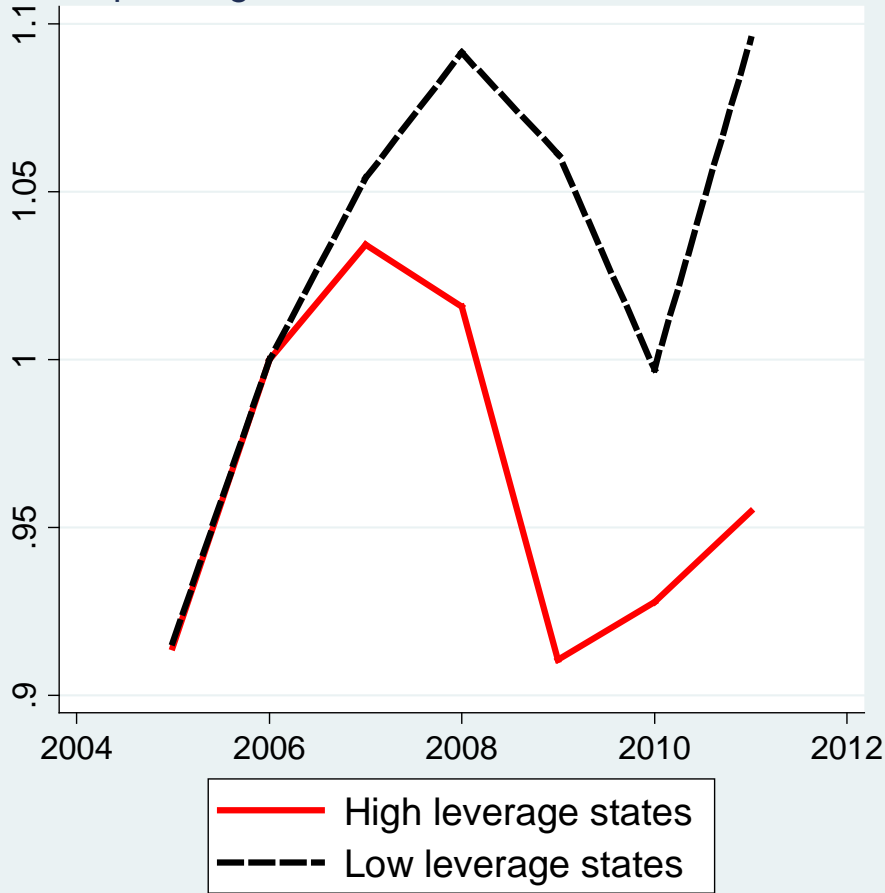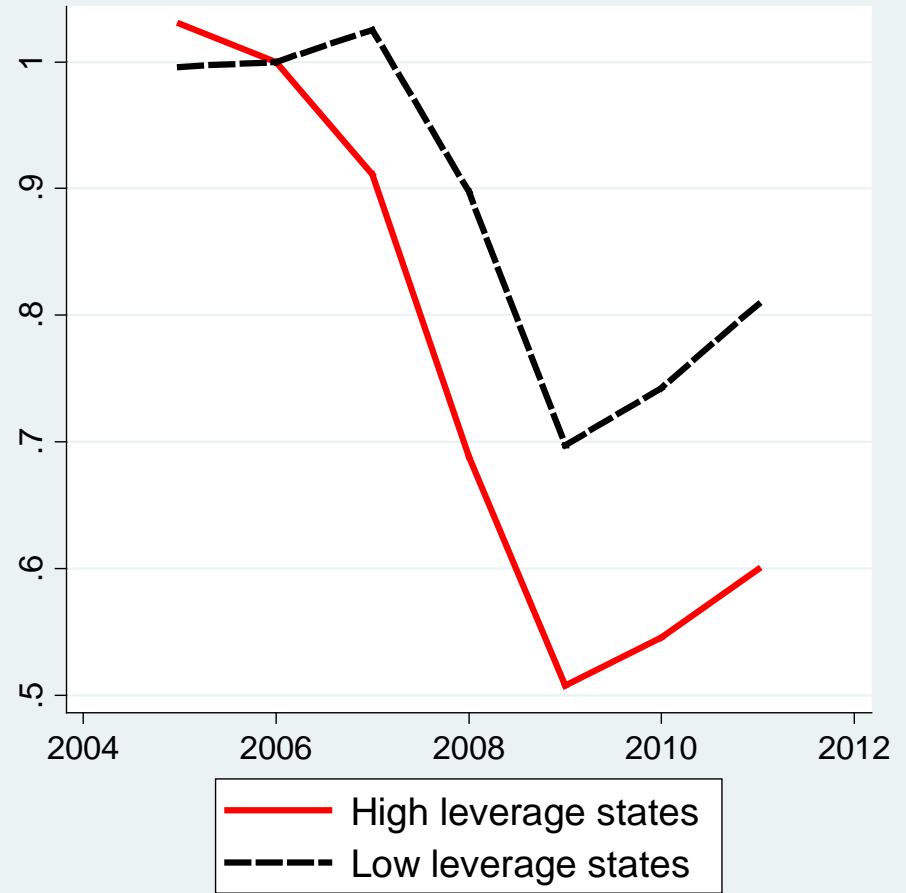

Employment to population ratio
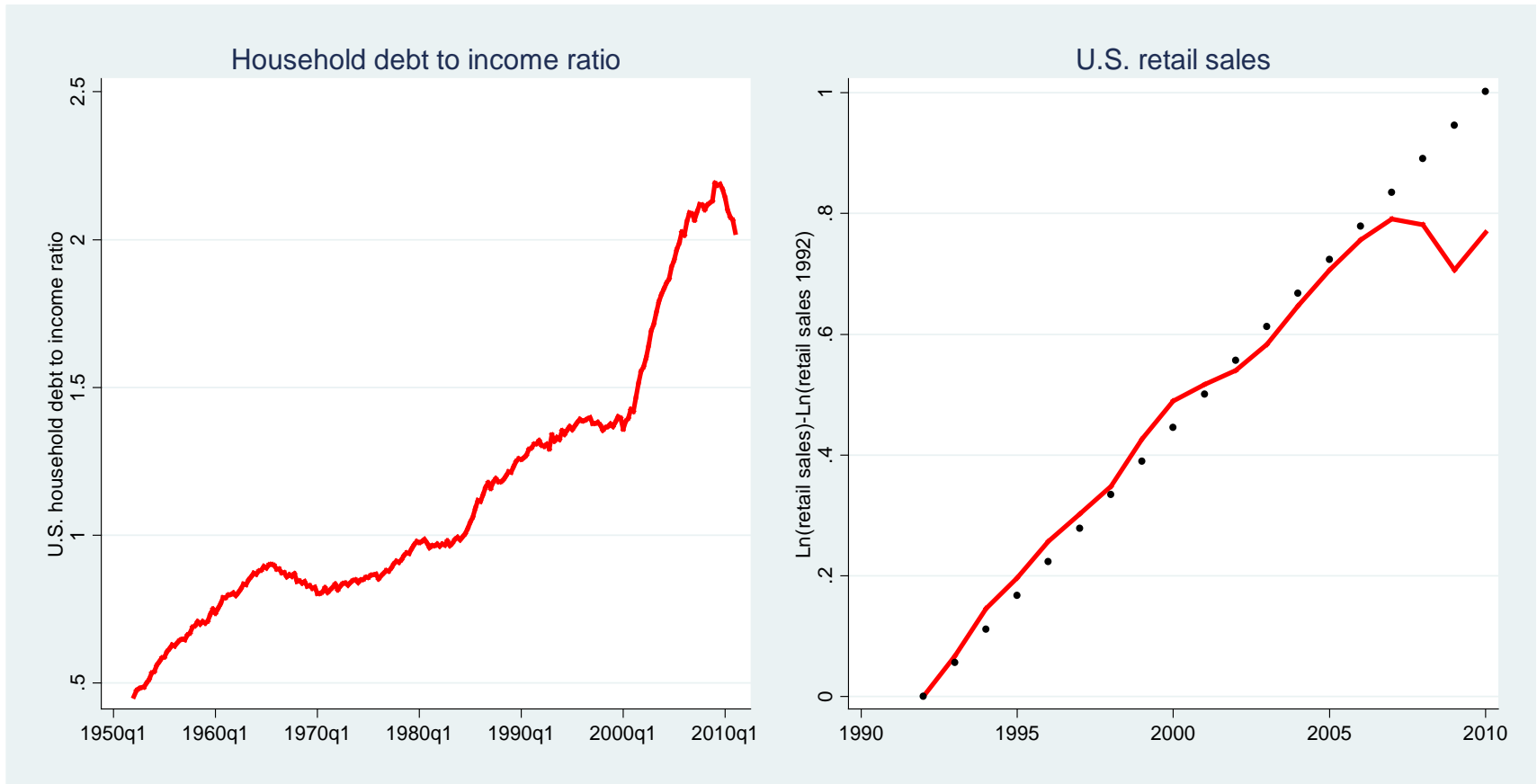
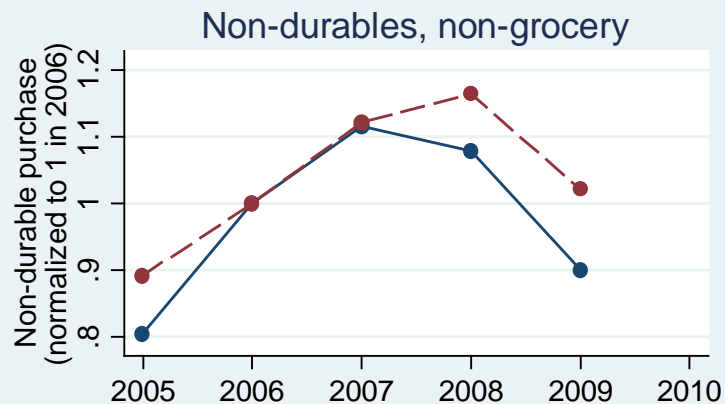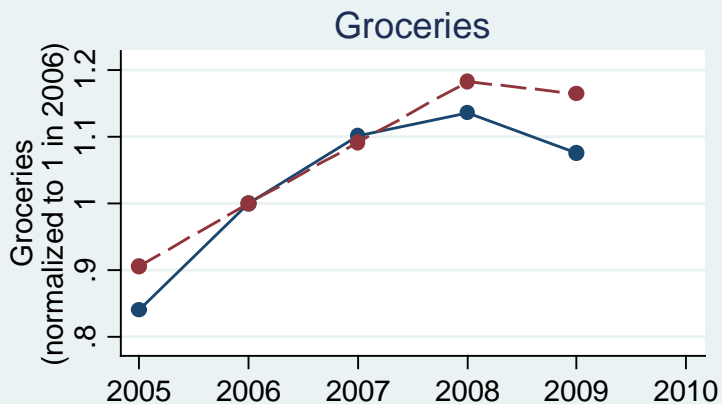# Heterogeneity in Wealth Shock in US
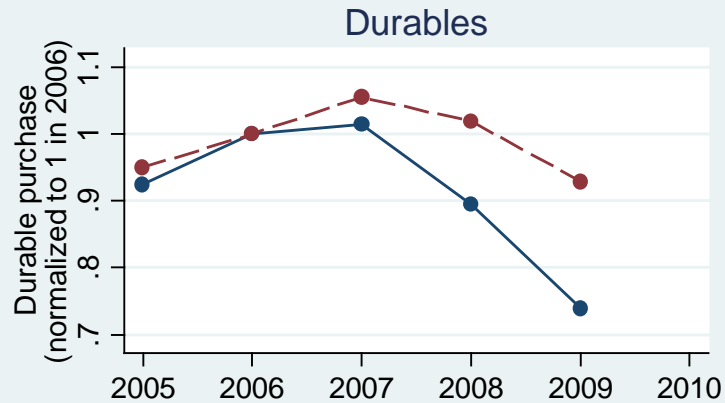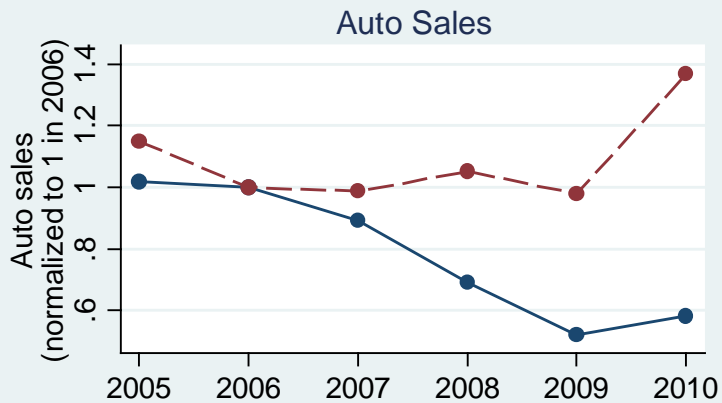
Using Geographic Variation in Household Leverage
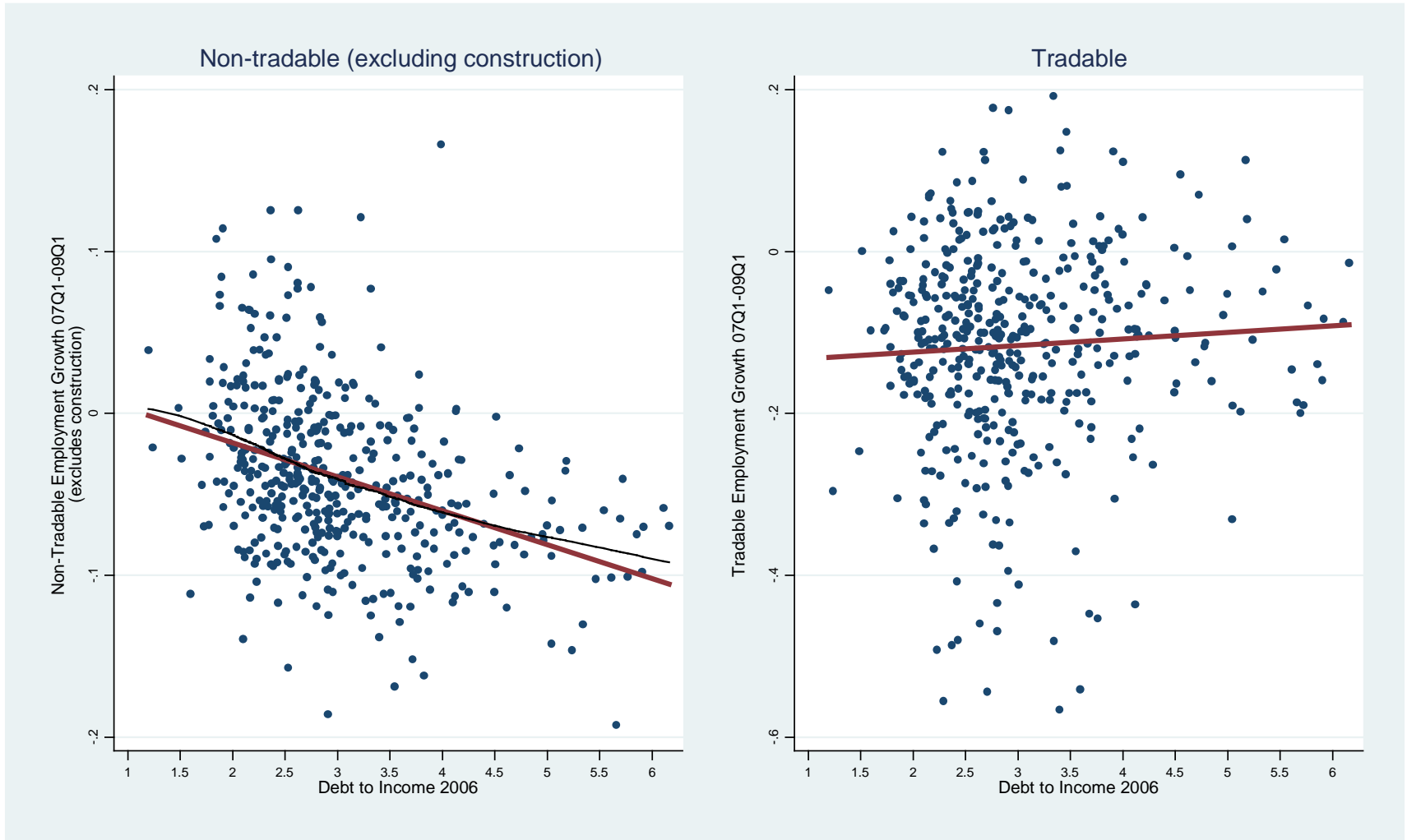

Spending, measured with sales tax revenue

Auto sales

Legend:
- High leverage states
- Low leverage states

Transparency ....

Exclusion Restrictions?

The GE Condition ...

4M Jobs Lost (65%)