

The Distribution of College Graduate Debt, 1990 to 2008: A Decomposition Approach

Brad J. Hershbein and Kevin M. Hollenbeck

W.E. Upjohn Institute for Employment Research

Upjohn/EPI/Spencer Conference on Student Loans
October 25, 2013
Ann Arbor, MI



**Education Policy
Initiative**
M | Ford School

- This project has two goals:
 - Describe how debt *distributions* have changed over time
 - Try to explain or decompose these changes
- Looking at distributions (not just means) is critical
 - Change in “tails” can affect mean, but leave most students unaffected
 - Right-hand tail is a different policy target than the “middle”
- Knowing role of “observables” also crucial
 - Changes in characteristics vs. changes in behavior
 - This dichotomy can vary across the distribution
 - Policies that understand this are likely to be more effective

- We focus on debt at graduation for bachelors recipients
 - Could also look at all students or other subgroups later on
- We use microdata to show cumulative debt distributions from 1990 through 2008 for all college grads and subsets of interest
- We employ multiple statistical decomposition techniques to parse out *which* factors caused *which* parts of the distribution to change *when*
 - Borrowing at all (Oaxaca-Blinder)
 - Entire distribution (DFL, RIF)

Key findings include:

- 1 Debt increased faster over the 1990s than the 2000s for grads
- 2 Increase in 2000s entirely in upper tail, at private schools, and due to private borrowing
- 3 Characteristics, including costs, explain about one-third of changes between 1990 and 2008
- 4 They generally explain more in the lower part of the distribution and less in the higher part
- 5 They also explain more between 1990 to 1996 and 2000 to 2008 than 1996 to 2000

- National Postsecondary Student Aid Survey (NPSAS)
 - Large, cross-sectional, nationally representative survey of students at Title IV institutions
 - Fielded approx. quadrennially: we use 1990, 1996, 2000, 2004, and 2008 waves
 - Specifically designed to collect info on how students pay for college
 - Has merged administrative data from FAFSA and NSLDS
 - Used as basis for longitudinal studies: Beginning Postsecondary Students and Baccalaureate and Beyond

Data Strengths and Limitations

- Strengths
 - Large sample sizes
 - Very rich financial aid data
 - Frequent availability
 - Allows analysis for subgroups of interest
- Limitations
 - Not longitudinal, can't look at repayment or debt-to-income
 - Attendance history not complete; only have current year
 - Asset and transfer data are limited
 - Most recent wave (2008) is before Great Recession

	1990	1996	2000	2004	2008
Sample Size	3,270	1,340	12,230	5,170	23,340
Weighted	724,000	897,000	1,217,000	1,448,000	1,822,000

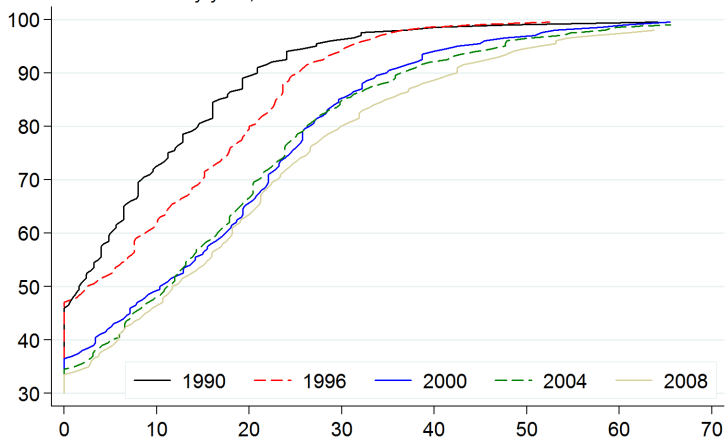
Cumulative borrowing statistics from NPSAS, by wave

	1990	1996	2000	2004	2008	2008r
<i>Ever borrow</i>	0.545	0.526	0.636	0.656	0.666	0.682
<i>Total borrowing (\$2012: 000s)</i>						
Mean	7.2	9.2	14.4	14.8	16.7	17.2
25 th	0.0	0.0	0.0	0.0	0.0	0.0
Median	1.9	2.5	10.9	11.6	12.1	13.1
75 th	11.4	17.7	24.5	23.8	26.6	26.6
90 th	20.8	25.4	34.8	36.4	42.5	42.5
95 th	27.3	30.8	42.5	47.7	51.6	52.1
99 th	48.1	44.9	60.6	65.6	85.0	85.0
<i>Total borrowing among borrowers (\$2012: 000s)</i>						
Mean	13.2	17.6	22.6	22.6	25.0	25.2
10 th	2.4	5.4	5.6	6.0	5.8	5.9
25 th	4.8	9.7	12.9	11.9	12.2	12.4
Median	10.4	17.0	21.8	20.4	21.3	21.3
75 th	18.0	23.6	29.3	29.8	33.0	33.1
90 th	25.7	30.2	38.8	42.6	47.8	47.8
95 th	32.1	35.1	49.0	51.6	56.2	56.0
99 th	64.2	51.6	64.5	72.7	90.3	90.3

Notes: Statistics use population weights and are for domestic students in the year indicated. Monetary amounts are inflated using the PCE index from the Bureau of Economic Analysis. Borrowing is from all sources except friends and family and excludes loans taken out by parents (PLUS loans).

Total Borrowing: All Graduates

Cumulative borrowing distribution among college graduates
by year, in thousands of constant 2012 dollars

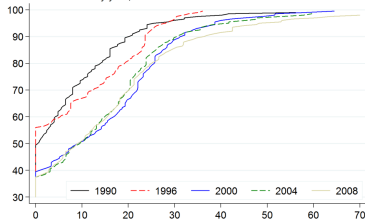


Sources: NPSAS, respective years

All calculations use sample weights and include student-level borrowing from all sources EXCEPT friends and family. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

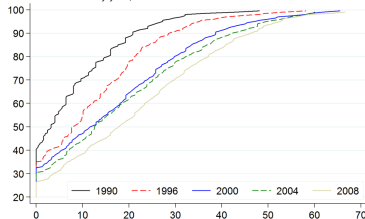
Total Borrowing: Subgroups 1

Cum. borrowing distribution among college graduates (dependents)
by year, in thousands of constant 2012 dollars



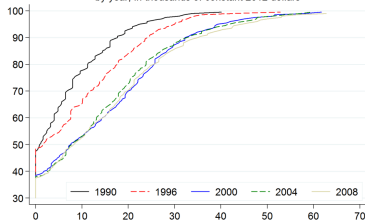
Sources: NPSAS, respective years
All calculations use sample weights and include student-level borrowing from all sources EXCEPT friends and family. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

Cum. borrowing distribution among college graduates (independents)
by year, in thousands of constant 2012 dollars



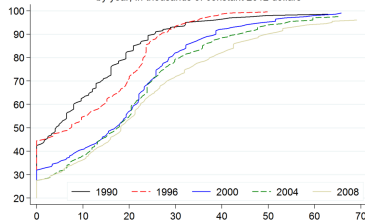
Sources: NPSAS, respective years
All calculations use sample weights and include student-level borrowing from all sources EXCEPT friends and family. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

Cum. borrowing distribution among college graduates (publics)
by year, in thousands of constant 2012 dollars



Sources: NPSAS, respective years
All calculations use sample weights and include student-level borrowing from all sources EXCEPT friends and family. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

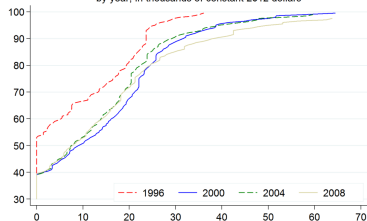
Cum. borrowing distribution among college graduates (private NFP)
by year, in thousands of constant 2012 dollars



Sources: NPSAS, respective years
All calculations use sample weights and include student-level borrowing from all sources EXCEPT friends and family. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

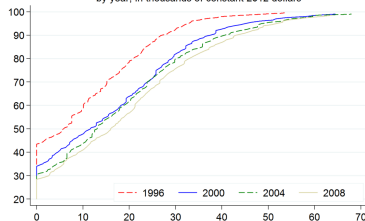
Total Borrowing: Subgroups 2

Cum. borrowing distribution among college graduates (≤ 4 years)
by year, in thousands of constant 2012 dollars



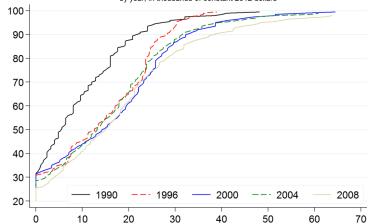
Sources: NPSAS, respective years
All calculations use sample weights and include student-level borrowing from all sources EXCEPT friends and family. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

Cum. borrowing distribution among college graduates (≥ 5 years)
by year, in thousands of constant 2012 dollars



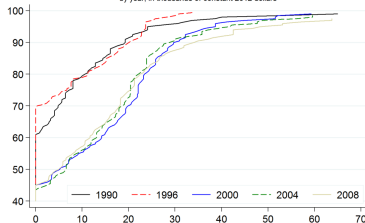
Sources: NPSAS, respective years
All calculations use sample weights and include student-level borrowing from all sources EXCEPT friends and family. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

Cum. borrowing dist. among college grads (dependents, bottom income half)
by year, in thousands of constant 2012 dollars



Sources: NPSAS, respective years
All calculations use sample weights and include student-level borrowing from all sources EXCEPT friends and family. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

Cum. borrowing dist. among college graduates (dependents, top income half)
by year, in thousands of constant 2012 dollars



Sources: NPSAS, respective years
All calculations use sample weights and include student-level borrowing from all sources EXCEPT friends and family. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

- Oaxaca-Blinder:

$$E[Y^B - Y^A] = E[X^B - X^A]\beta^A + E[X^B][\beta^B - \beta^A]$$

Decomposition Techniques

- Oaxaca-Blinder:

$$E[Y^B - Y^A] = E[X^B - X^A]\beta^A + E[X^B][\beta^B - \beta^A]$$

- Semiparametric reweighting (DiNardo, Fortin, and Lemieux 1996)
 - Reweight data on observables from group B to resemble joint distribution of $\mathbf{X}$ from group A
 - Creates counterfactual distribution and more robust to functional form violations than O-B
 - However, hard to identify role of specific X ; not path invariant

Decomposition Techniques

- Oaxaca-Blinder:

$$E[Y^B - Y^A] = E[X^B - X^A]\beta^A + E[X^B][\beta^B - \beta^A]$$

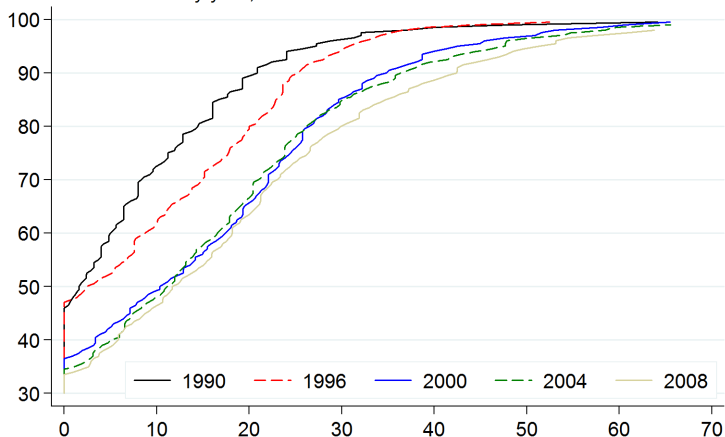
- Semiparametric reweighting (DiNardo, Fortin, and Lemieux 1996)
 - Reweight data on observables from group B to resemble joint distribution of $\mathbf{X}$ from group A
 - Creates counterfactual distribution and more robust to functional form violations than O-B
 - However, hard to identify role of specific X ; not path invariant
- Recentered influence functions (Firpo, Fortin, Lemieux 2007)
 - $RIF_q = Y_q + \frac{q}{f(Y_q)} - \frac{1}{f(Y_q)} \cdot \mathbf{1}(Y \leq Y_q)$
 - q is a quantile, $f(Y_q)$ is an (estimated) density at q , $E[RIF_q] = Y_q$
 - By running O-B on RIF_q , get decomp at unconditional quantiles

Decomposition Techniques: the Xs

- Age (dummies), dependency, gender, ethnicity, marital status, state of residence, region of school, in-state student, parental education, full vs. part-time, full vs. part-year, changed schools dummy, majors, sector of school, quartic in EFC by dependency, quartic in list tuition (cost of attendance), quartic in grants, full interactions of costs and grants
- Explicit decision **not** to use quartic in *net cost*
 - It would imply restrictions on coefficients of flexible interactions
 - The data soundly reject these restrictions
- All variables are measured during the final year of attendance before graduation

Total Borrowing: All Graduates

Cumulative borrowing distribution among college graduates
by year, in thousands of constant 2012 dollars



Sources: NPSAS, respective years

All calculations use sample weights and include student-level borrowing from all sources EXCEPT friends and family. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

O-B: Ever Borrow

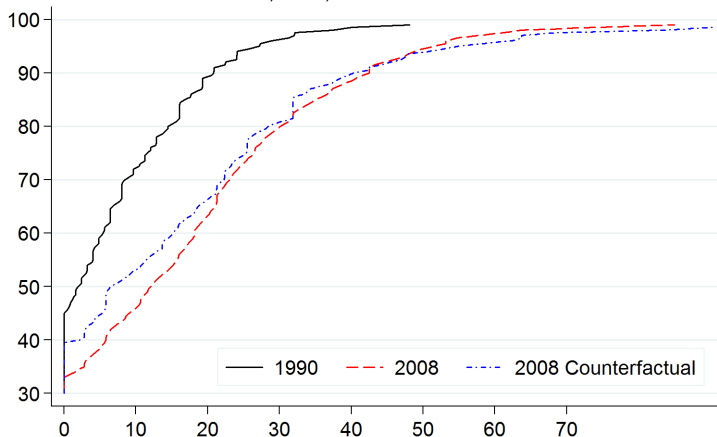
Oaxaca-Blinder Decompositions of Ever Borrowed

	2008–1990		2000–1996	
Mean difference (pp)	12.04	(1.43)	10.99	(2.02)
<i>Composition effects due to:</i>				
Age/dependency status	-0.87	(0.45)	-0.09	(0.39)
Sex, marital status, ethnicity	0.99	(0.55)	0.85	(0.52)
Parental education	-0.50	(0.46)	0.98	(0.69)
Location, in-state status	-1.80	(0.65)	0.41	(0.83)
School sector, attendance, major	0.58	(1.06)	1.75	(0.69)
Expected family contribution	0.55	(0.42)	-1.08	(0.84)
Tuition and grants	6.44	(1.95)	-0.73	(1.33)
Total	5.38	(2.36)	2.07	(2.16)
<i>Structural effects due to:</i>				
Age/dependency status	5.92	(1.62)	-0.08	(2.54)
Sex, marital status, ethnicity	-5.16	(4.32)	11.94	(5.59)
Parental education	0.33	(1.09)	1.13	(1.97)
Location, in-state status	1.49	(2.16)	-1.77	(2.95)
School sector, attendance, major	-2.83	(10.65)	-5.92	(6.25)
Expected family contribution	-1.85	(2.16)	9.90	(3.53)
Tuition and grants	-2.50	(6.20)	-4.20	(8.18)
Constant	11.24	(14.40)	-2.07	(11.54)
Total	6.66	(2.22)	8.93	(2.02)

Notes: Each column refers to the later period less the earlier period. Oaxaca-Blinder decompositions are based on coefficients from the base period reference and are estimated via OLS (with sample weights). Standard errors robust to heteroskedasticity and intra-college correlation are in parentheses. Borrowing is from all sources except friends and family and excludes loans taken out by parents (PLUS loans). Results change trivially if time to degree is included for the latter two panels.

DFL 1990 to 2008

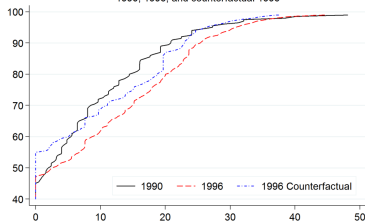
CDF of borrowing among college graduates in 000s of \$2012
1990, 2008, and counterfactual 2008



The DFL (1996) reweighting procedure is used to create a counterfactual distribution for 2008, assuming the distribution of covariates was the same as in 1990. See text for set of covariates. Student-level borrowing is from all sources except friends and family. Sources: NPSAS, respective years.

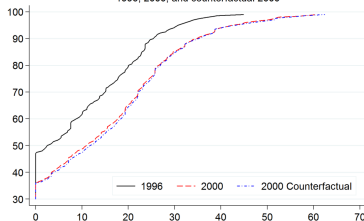
DFL: Other years

CDF of borrowing among college graduates in 000s of \$2012
1990, 1996, and counterfactual 1996



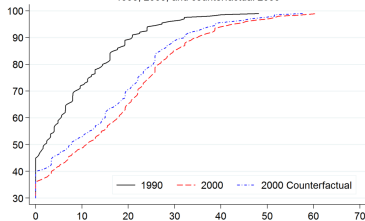
The DFL (1996) reweighting procedure is used to create a counterfactual distribution for 1996, assuming the distribution of covariates was the same as in 1990. See text for set of covariates. Student-level borrowing is from all sources except friends and family. Sources: NPSAS, respective years.

CDF of borrowing among college graduates in 000s of \$2012
1996, 2000, and counterfactual 2000



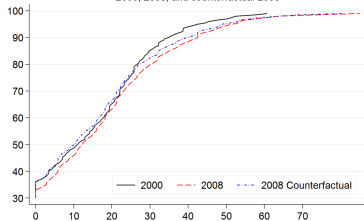
The DFL (1996) reweighting procedure is used to create a counterfactual distribution for 2000, assuming the distribution of covariates was the same as in 1996. See text for set of covariates. Student-level borrowing is from all sources except friends and family. Sources: NPSAS, respective years.

CDF of borrowing among college graduates in 000s of \$2012
1990, 2000, and counterfactual 2000



The DFL (1996) reweighting procedure is used to create a counterfactual distribution for 2000, assuming the distribution of covariates was the same as in 1990. See text for set of covariates. Student-level borrowing is from all sources except friends and family. Sources: NPSAS, respective years.

CDF of borrowing among college graduates in 000s of \$2012
2000, 2008, and counterfactual 2008



The DFL (2000) reweighting procedure is used to create a counterfactual distribution for 2008, assuming the distribution of covariates was the same as in 2000. See text for set of covariates. Student-level borrowing is from all sources except friends and family. Sources: NPSAS, respective years.

RIF Decompositions of Borrowing: 2008–1990

	Mean	50 th percentile	75 th percentile	90 th percentile
Difference (000s \$2012)	9.44	10.21	15.16	21.66
<i>Composition effects due to:</i>				
Age/dependency status	-0.19	0.42	0.54	0.21
Sex, marital status, ethnicity	0.12	0.18	0.14	-0.11
Parental education	0.00	0.08	0.27	0.14
Location, in-state status	-0.16	-0.37	-0.18	0.00
School sector, attendance, major	0.57	0.96	0.96	1.65
Expected family contribution	0.24	0.67	1.08	1.28
Tuition and grants	1.16	1.32	2.67	-4.06
<i>Total</i>	1.74	3.26	5.48	-0.89
<i>Structural effects due to:</i>				
Age/dependency status	-0.16	0.94	-1.46	-3.05
Sex, marital status, ethnicity	-1.23	-2.48	-2.92	-4.05
Parental education	0.02	0.23	-0.75	-0.32
Location, in-state status	-0.41	-0.58	-0.11	-0.34
School sector, attendance, major	0.28	-2.82	1.11	-1.55
Expected family contribution	-0.29	-1.40	0.81	2.93
Tuition and grants	4.26	8.16	3.40	9.64
Constant	5.23	4.89	9.60	16.20
<i>Total</i>	7.70	6.94	9.68	22.55

Notes: Each column refers to the later period less the earlier period. The recentered influence functions and quantiles are calculated with sample weights; the decompositions are based on coefficients from the base period reference and are estimated via OLS (without sample weights). Inference is based on bootstrapped standard errors (100 replications). Borrowing is from all sources except friends and family and excludes loans taken out by parents (PLUS loans).

Decomposition Summary

- From 1990 to 2008, techniques produce similar results:
 - Observables explain about half at borrowing margin
 - They explain between $1/3$ and $1/2$ at median, less above
 - Costs alone are half of explained share

Decomposition Summary

- From 1990 to 2008, techniques produce similar results:
 - Observables explain about half at borrowing margin
 - They explain between $1/3$ and $1/2$ at median, less above
 - Costs alone are half of explained share
- For other time periods:
 - For DFL, observables explain between half and all of change in early 1990s and over 2000s
 - RIFs are similar, but overexplain change during 2000s
 - Increasing costs more important in 2000s than 1990s

Decomposition Summary

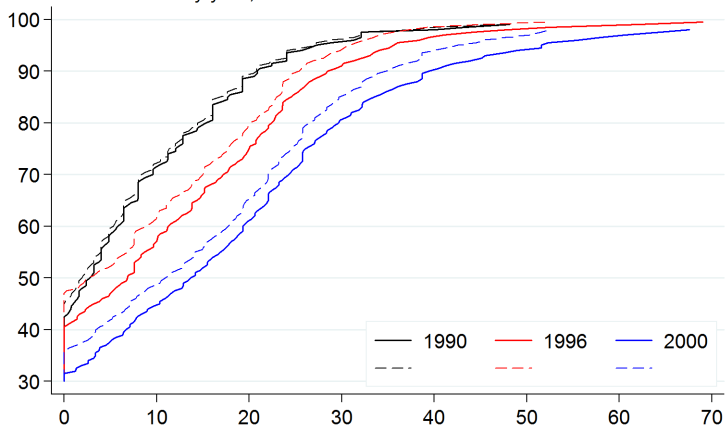
- From 1990 to 2008, techniques produce similar results:
 - Observables explain about half at borrowing margin
 - They explain between $1/3$ and $1/2$ at median, less above
 - Costs alone are half of explained share
- For other time periods:
 - For DFL, observables explain between half and all of change in early 1990s and over 2000s
 - RIFs are similar, but overexplain change during 2000s
 - Increasing costs more important in 2000s than 1990s
- Observables explain almost nothing between 1996 and 2000
 - This period is when debt grew fastest
 - Role for unobservables suggests policy changes

Other Possibilities

- Formal loans have replaced informal loans?
 - “Informal” loans are from friends and family; no credit reporting
- Parents are transferring burden to their children
 - Student-level loans replacing parent-level PLUS loans
- Interest rate changes
- Unsubsidized loans
- Growing availability of non-federal loans
 - Under optimal behavior, should matter in upper tail

Informal loans

CDF of college graduate borrowing, with informal loans
by year, in thousands of constant 2012 dollars



Sources: NPSAS, respective years

All calculations use sample weights. Solid lines include student-level borrowing from all sources INCLUDING friends and family.

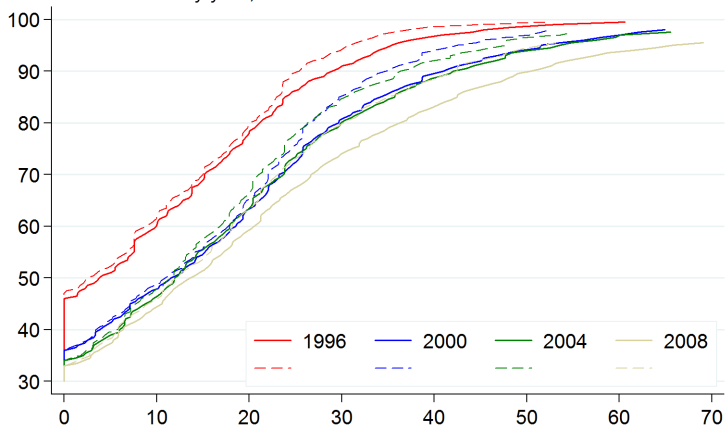
Dashed lines EXCLUDE loans from friends and family. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

Other Possibilities

- Formal loans have replaced informal loans?
 - “Informal” loans are from friends and family; no credit reporting
- Parents are transferring burden to their children
 - Student-level loans replacing parent-level PLUS loans
- Interest rate changes
- Unsubsidized loans
- Growing availability of non-federal loans
 - Under optimal behavior, should matter in upper tail

PLUS loans

CDF of college graduate borrowing, with PLUS loans
by year, in thousands of constant 2012 dollars



Sources: NPSAS, respective years

All calculations use sample weights. Solid lines include student-level borrowing from all sources INCLUDING parental PLUS loans.

Dashed lines EXCLUDE parental PLUS loans. Borrowed amounts are self-reported in 1990; for other years, the federal component

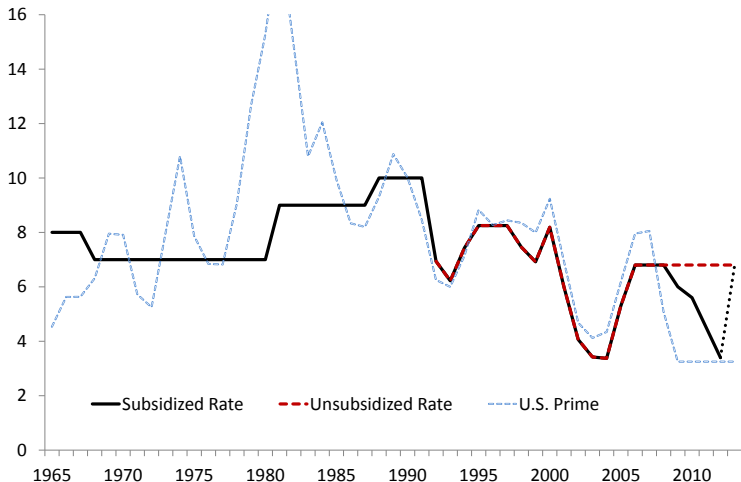
is from administrative data and other borrowing is self-reported.

Other Possibilities

- Formal loans have replaced informal loans?
 - “Informal” loans are from friends and family; no credit reporting
- Parents are transferring burden to their children
 - Student-level loans replacing parent-level PLUS loans
- Interest rate changes
- Unsubsidized loans
- Growing availability of non-federal loans
 - Under optimal behavior, should matter in upper tail

Interest rates

Interest Rates on New Undergraduate Stafford Loans, 1965–2013



Sources: U.S. Senate Budget Bulletin, August 4, 2006; <http://www.finaid.org/loans/historicalrates.phtml>

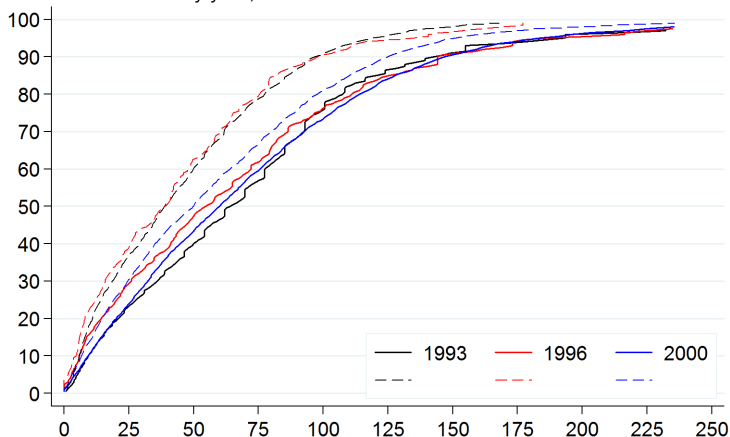
Note: All federal loans were subsidized until 1992, when unsubsidized loans became available.

Other Possibilities

- Formal loans have replaced informal loans?
 - “Informal” loans are from friends and family; no credit reporting
- Parents are transferring burden to their children
 - Student-level loans replacing parent-level PLUS loans
- Interest rate changes
- Unsubsidized loans
- Growing availability of non-federal loans
 - Under optimal behavior, should matter in upper tail

Unsubsidized loans

CDF of family income, All and Stafford borrowers
by year, in thousands of constant 2012 dollars

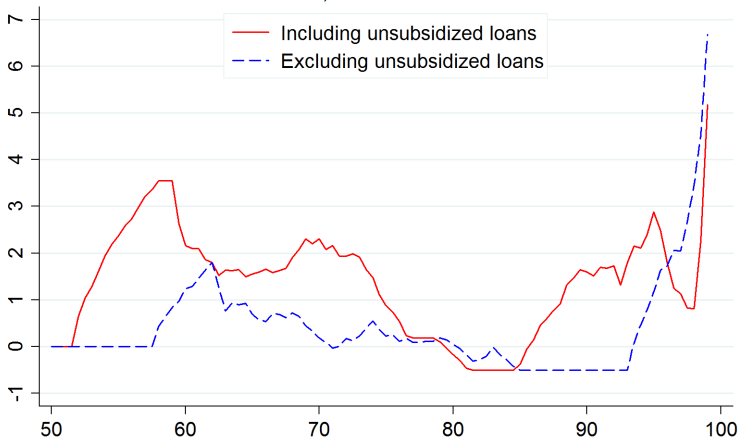


Sources: NPSAS, respective years

All calculations use sample weights. Solid lines include all graduates; dashed lines include ONLY students who took out a Stafford loan their senior year.

Unsubsidized loans

Difference in senior year borrowing CDFs, w/ and w/o unsubsidized loans
Between 2000 and 1996, in thousands of constant 2012 dollars



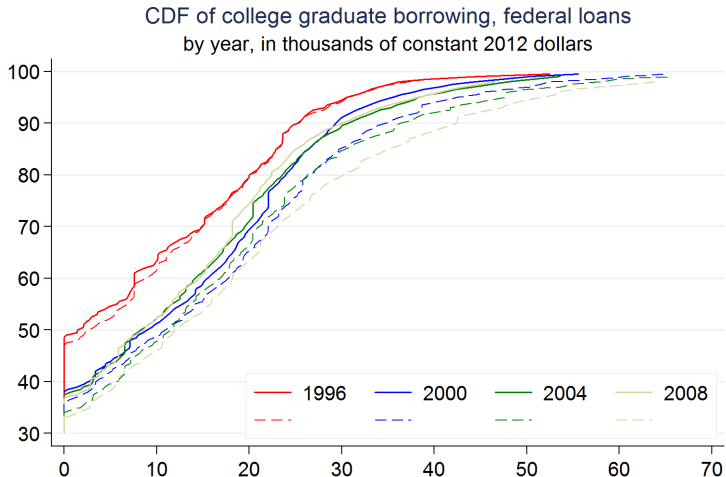
Sources: NPSAS, respective years

All calculations use sample weights. Solid lines include all formal student-level borrowing in senior year. Dashed lines exclude unsubsidized Stafford loans. The federal component of borrowing is from administrative data and other borrowing is self-reported.

Other Possibilities

- Formal loans have replaced informal loans?
 - “Informal” loans are from friends and family; no credit reporting
- Parents are transferring burden to their children
 - Student-level loans replacing parent-level PLUS loans
- Interest rate changes
- Unsubsidized loans
- Growing availability of non-federal loans
 - Under optimal behavior, should matter in upper tail

Nonfederal loans



Sources: NPSAS, respective years

All calculations use sample weights. Solid lines include student-level borrowing from federal loans only. Dashed lines include student-level borrowing as in Figure 1. Borrowed amounts are self-reported in 1990; for other years, the federal component is from administrative data and other borrowing is self-reported.

Summary

- Debt profiles increased more in 1990s than 2000s
 - Costs and EFC explain about half 1990–1996
 - 1996–2000 increase likely due to unsubsidized and private loan availability

Summary

- Debt profiles increased more in 1990s than 2000s
 - Costs and EFC explain about half 1990–1996
 - 1996–2000 increase likely due to unsubsidized and private loan availability
- Recent increase has been mostly in upper tail
 - Almost entirely due to nonfederal loans
 - Changes in observables explain nearly all of increase
 - Costs drive the bulk, but other factors matter, too

Summary

- Debt profiles increased more in 1990s than 2000s
 - Costs and EFC explain about half 1990–1996
 - 1996–2000 increase likely due to unsubsidized and private loan availability
- Recent increase has been mostly in upper tail
 - Almost entirely due to nonfederal loans
 - Changes in observables explain nearly all of increase
 - Costs drive the bulk, but other factors matter, too
- Overall from 1990 to 2008:
 - Observables explain between $1/3$ to $1/2$ of increase in the middle...
 - ... and 0 to $1/4$ at the top
 - Costs alone are about half of explained share

Conclusions and Caveats

- Data run only through 2008, before Great Recession
 - Private loan volume fell a lot, but *slowly* recovering
 - “Average debt” still increased, but not by much
 - Full NPSAS 2012 data come out next year...

Conclusions and Caveats

- Data run only through 2008, before Great Recession
 - Private loan volume fell a lot, but *slowly* recovering
 - “Average debt” still increased, but not by much
 - Full NPSAS 2012 data come out next year...
- Debt is growing relatively slowly for *most* graduates
 - What are characteristics of students in top decile?

Conclusions and Caveats

- Data run only through 2008, before Great Recession
 - Private loan volume fell a lot, but *slowly* recovering
 - “Average debt” still increased, but not by much
 - Full NPSAS 2012 data come out next year...
- Debt is growing relatively slowly for *most* graduates
 - What are characteristics of students in top decile?
- Supply-side factors may increase debt more than demand-side factors, especially higher up in distribution

Conclusions and Caveats

- Data run only through 2008, before Great Recession
 - Private loan volume fell a lot, but *slowly* recovering
 - “Average debt” still increased, but not by much
 - Full NPSAS 2012 data come out next year...
- Debt is growing relatively slowly for *most* graduates
 - What are characteristics of students in top decile?
- Supply-side factors may increase debt more than demand-side factors, especially higher up in distribution
- But need more research on how institutions “capture” financial aid and how this relates to debt and student success