Examining Income Expectations in the College and Early Post-college Periods: New Distributional Tests of Rational Expectations

Thomas Crossley, Yifan Gong, Todd Stinebrickner, Ralph Stinebrickner

University of Michigan, University of Nebraska-Lincoln, University of Western Ontario, Berea College

Teaching Materials

Rationality of Income Beliefs

- Many decisions made by young adults are influenced by their beliefs about future income.
  - e.g., college attendance, consumption-saving allocation (Sandmo, 1970; Leland, 1978)
- Optimality of these decisions depends on rationality of income beliefs.
  - Rational Expectations (RE): Beliefs can be characterized by the conditional distribution of the outcome given an individual's information set. (Muth, 1961)
- Most existing related research:
  1. test implications of RE for *mean-beliefs* elicited using **point expectations questions** (see Pesaran and Weale, 2006, for a survey);
     - Decisions often also depend on higher moments of belief distributions, e.g. precautionary savings and income uncertainty.
  2. focus on post-college income beliefs.
     - In-school beliefs matter for college attendance/dropout decisions.

Our Contribution

1. We develop several new distributional tests of RE that exploit implications for individual *belief distributions* elicited using **probabilistic expectations questions**.
   - probabilistic expectations questions: e.g. perceived probabilities for fixed outcome bins, fixed percentiles of belief distributions
   - Asymptotically, our tests can detect more types of violations of RE (e.g., incorrect amount of uncertainty) than existing mean-based RE tests (e.g., Regression Test).

2. We apply our tests to longitudinal expectations data covering both in-school and post-college periods from the Berea Panel Study.
   - RE provides a better description of income beliefs for the post-college period than for the in-school period.
   - Our new tests reject more often than existing mean-based RE tests in practice.

Literature Review

- Tests of Rational Expectations and applications to income expectations.
  - Pesaran and Weale (2006) provides a survey of the literature.
  - Dominitz (1998) and D'Haultfoeuille, Gaillac, and Maurel (2021) (DGM, henceforth) are most relevant to this paper.
- A growing literature on students' beliefs and expectations.
  - e.g., Zafar (2011); Arcidiacono, Hotz, and Kang (2012); Stinebrickner and Stinebrickner (2012)
- Papers studying survey data on probabilistic expectations about income.
  - e.g., Guiso, Jappelli, and Terlizzese (1992); Dominitz and Manski (1996, 1997a,b); Das and Van Soest (1997); Das and Donkers (1999); Das, Dominitz, and Van Soest (1999)

Rational Expectations Hypothesis

- Outcome Y_i depends on i.i.d. factors X_i.
  - We can allow Y_i to depend on aggregate factors Z.
- At an early time t, individual i observes a subset of X_i, $X_i^{\theta_i, t-}$.
  - Random variable Θ_i determines which factors are observed at t.
- Y_{it}^B describes beliefs about Y_i at t.

Rational Expectations

Individual i has Rational Expectations about Y_i if and only if $Y_{it}^B \overset{d}{=} Y_i | (X_i^{\theta_i, t-} = x_i^{\theta_i, t-}, \Theta_i = \theta_i)$.

Data Environment

- We observe realization y_i and partial information about belief distribution Y_{it}^B in an unbalanced panel.
  - balanced panel and repeated cross-sections as special cases
  - assume MCAR for main results - can relax

- Information about Y_{it}^B is elicited using **probabilistic expectations** questions.
  - *Probability Format*: perceived probability $P_{it}^{B,k} \equiv Prob(Y_{it}^B \in S_k)$ for outcome bin S_k, $k = 1, 2, ..., K$ Question
  - *Percentile Format*: r_j-th percentile (e.g., minimum, maximum, quartiles), $C_{it}^{r_j,B}$, of Y_{it}^B, $j = 1, 2, ..., J$ Question

- Under parametric assumptions, we can compute moments (e.g., mean μ_{it}^B, variance v_{it}^B) of Y_{it}^B from reported probabilistic expectations.

Tests based on Marginal Distributions of Y_{it}^B and y_i

- When data come from repeated cross-sections, we only have information about marginal distributions of Y_{it}^B and y_i.
  - Y^R denotes the distribution of y_i.
  - $Y_t^{AB} \equiv E_i(Y_{it}^B)$: Aggregate Belief Distribution
- The only implication of RE in this case is $Y_t^{AB} \overset{\text{d}}{=} Y^R$.
- Two Corollaries:

CDF-Based Aggregate-Distribution Test

$$E_i(P_{it}^{B,k}) = Prob(Y^R \in S_k), \text{ for all } k.$$

Moment-Based Aggregate-Distribution Test

$$E_i(\mu_{it}^B) = E_i(y_i) \text{ and } E_i(v_{it}^B) + var_i(\mu_{it}^B) = var_i(y_i).$$

Comparing CDF-Based and Moment-Based Aggregate-Distribution Tests

- CDF-Based Aggregate-Distribution Test:
  - nonparametric
  - robust to classical reporting error
  - exploits all implications of RE for marginal distributions when data elicited using the probability format
- Moment-Based Aggregate-Distribution Test:
  - implementable with any format of probabilistic expectations data
  - individual components are easier to interpret
  - easily adaptable to accommodate additive aggregate shocks

Tests based on Joint Distribution of Y_{it}^B and y_i

- When data come from (unbalanced) panels, we have information about the joint distribution of Y_{it}^B and y_i.
- RE has many more implications in this case.
  - We consider two intuitive and easy-to-implement tests in this paper.

Uncertainty Test

$$E_i(v_{it}^B) = var_i(y_i - \mu_{it}^B).$$

Percentile-Based Joint-Distribution Test (Dominitz, 1998)

$$E_i[I(y_i \leq C_{it}^{r,B})] = 0.01r \text{ for any } r \in [0, 100],$$

where $C_{it}^{r,B}$ is the r-th percentile of Y_{it}^B.

Relationship with Existing Mean-Based RE Tests

- Marginal Distribution Mean-based Test - DGM Test
  - Y^R is a mean preserving spread of the distribution of μ_{it}^B.
- When Y_{it}^B is fully observed, CDF-based Aggregate-Distribution Test $\rightarrow$ Mean-based DGM Test.
  - complementary in practice because information about Y_{it}^B is partial
- Joint Distribution Mean-based Test - Regression Test
  - regress y_i on μ_{it}^B; slope $= 1$, intercept $= 0$
  - test 1) no mean-bias and 2) correct covariance between mean-beliefs and realizations
- Moment-based Aggregate Distribution Test + Uncertainty Test $\rightarrow$ Regression Test
  - test 3) correct amount of uncertainty in addition to 1) and 2)

Berea Panel Study and Berea College

- The Berea Panel Study (BPS) is a longitudinal case study that follows two cohorts of students at Berea College from the time of entrance in 2000 and 2001, until 2014.
- Berea College is a liberal art college in central Kentucky.
  - Similar to typical nearby four-year institutions along many dimensions, e.g., ACT score, spending on amenities/academic, etc.
- Students at Berea College are mostly coming from low-income backgrounds.
  - Likely to have information problem.

Data Structure

- In-School Beliefs:
  - about own income at age 28
  - elicited using the percentile format (min, max and three quartiles)
  - elicited at the time of entrance and end of each academic year
- Post-College Beliefs:
  - about family income one year ahead and five years ahead
  - elicited using the probability format (ten bins, each covering \$15,000)
  - elicited around the middle of each post-college year
- Income Realization:
  - both own and spousal income (if present)
  - elicited around the middle of each post-college year
- When needed, we compute μ_{it}^B and v_{it}^B under a step-wise uniform distributional assumption; all values deflated to 2001 dollars using CPI.

Compare In-School Beliefs and Post-College Beliefs

- Compare in-school beliefs about own income at age 28 and post-college beliefs about family income at age 28.
- In-school beliefs are directly available.
  - 10-year ahead (time of entrance) to 6-year ahead (time of graduation)
- Post-college beliefs can be constructed as 1-year ahead beliefs for 27 years old and 5-year-ahead beliefs for 23 years old.

Descriptive Statistics for In-School and Post-College Beliefs

Table 1: Descriptive Statistics: In-School and Post-College Beliefs about Income at Age 28

Item	$E_i(\mu_{it}^B)$	$E_i(y_i)$	$E_i[v_{it}^B] + var_i(\mu_{it}^B)$	$var_i(y_i)$	$E_i[v_{it}^B]$	$var_i(y_i - \mu_{it}^B)$
Panel A: In-School (Own Income at 28)						
10-year ahead	46.2	28.2	456.4	169.0	123.6	442.9
9-year ahead	45.6	28.2	383.8	165.5	109.8	333.4
8-year ahead	41.7	28.7	305.3	172.1	101.4	310.8
7-year ahead	39.5	29.0	259.5	168.8	87.1	271.0
6-year ahead	33.8	27.9	192.4	163.4	72.3	178.3
Panel B: Post-College (Family Income at 28)						
5-year ahead	45.8	42.8	468.7	543.7	162.3	432.1
1-year ahead	41.9	42.2	454.0	514.3	117.1	180.4

Notes: The unit of annual income is $1,000 USD. Panel A reports the objects relevant for testing the rationality of students' in-school beliefs about own annual income at age 28. Panel B reports the objects relevant for testing the rationality of students' post-college beliefs about family annual income at age 28.

Testing Results - Parametric Tests

Table 2: Parametric Tests for Beliefs about Income at Age 28 (p-values)

10-year ahead	9-year ahead	8-year ahead	7-year ahead	6-year ahead	5-year ahead	1-year ahead
Mean						
$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	0.0762	0.7077
Variance						
$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	**0.0001**	**0.0039**	0.2079	0.3030	0.0797
Uncertainty						
$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	0.0687
Joint						
$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	**0.0001**	0.2171

- almost always reject (at a 5% level) for in-school beliefs
- rarely reject (at a 5% level) for post-college beliefs

Testing Results - Other Tests

Horizon	10-year ahead	9-year ahead	8-year ahead	7-year ahead	6-year ahead
Panel A: Fraction of Realizations below Certain Percentiles					
25th	0.6863	0.7181	0.6359	0.5990	0.4928
50th	0.8192	0.8340	0.7913	0.7525	0.6522
75th	0.8930	0.8996	0.8495	0.8366	0.7778
Panel B: p-value					
Common:	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$
Sample Size: n^{BY}	271	259	206	202	207

- The nonparametric Percentile-based Joint-Distribution Test rejects for all in-school beliefs as well.
- Mean-based Tests (DGM, Regression) produce similar results - almost always reject for in-school beliefs and rarely reject for post-college beliefs.

Comparisons with Mean-based Tests

- Asymptotically, our new distributional tests can detect more violations of RE than existing mean-based tests do.
  - The null of the CDF-based Aggregate Distribution Test is stronger than the mean-based DGM Test when there are infinitely many outcome bins.
  - The joint null of Moment-based Aggregate Distribution Test and Uncertainty Test is stronger than the Regression Test.
- We examine whether this is the case for post-college beliefs in finite sample.

CDF-Based Aggregate Distribution and DGM Tests

Table 3: CDF-Based Aggregate-Distribution Test about Income in 2007-2013 (p-values)

2007	2008	2009	2010	2011	2012	2013
1-year ahead Beliefs						
0.0002	0.1498	$< \mathbf{10^{-4}}$	0.3691	0.3235	**0.0015**	0.0914
5-year ahead Beliefs						
N.A.	N.A.	N.A.	N.A.	0.0808	**0.0167**	**0.0010**

- The CDF-Based Aggregate Distribution rejects for several year-horizon combinations.
- The mean-based DGM test does not reject for all post-college years.
  - p-value > 0.1 for all years

Parametric Tests and Regression Test

Year	2007	2008	2009	2010	2011	2012	2013
1-year ahead Beliefs							
Regression	**0.0002**	**0.0004**	**0.0213**	0.5514	0.1021	**0.0002**	**0.0001**
Mean	0.4448	0.2299	0.3736	0.4750	0.3902	**0.0000**	**0.0056**
Variance	**0.0137**	0.3669	0.5582	0.2496	0.6849	0.1129	**0.0060**
Uncertainty	0.0730	**0.0055**	**0.0059**	**0.0459**	**0.0303**	**0.0009**	**0.0028**
Joint	**0.0007**	**0.0223**	**0.0364**	0.2676	0.1926	$< \mathbf{10^{-4}}$	**0.0114**
5-year ahead Beliefs							
Regression	N.A.	N.A.	N.A.	N.A.	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	0.2981
Mean	N.A.	N.A.	N.A.	N.A.	**0.0134**	0.2218	0.7128
Variance	N.A.	N.A.	N.A.	N.A.	0.8124	0.2787	**0.0198**
Uncertainty	N.A.	N.A.	N.A.	N.A.	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	**0.0002**
Joint	N.A.	N.A.	N.A.	N.A.	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$

Parametric Tests and Regression Test

Year	2007	2008	2009	2010	2011	2012	2013
1-year ahead Beliefs							
Regression	**0.0002**	**0.0004**	**0.0213**	0.5514	0.1021	**0.0002**	**0.0001**
Mean	0.4448	0.2299	0.3736	0.4750	0.3902	**0.0000**	**0.0056**
Variance	**0.0137**	0.3669	0.5582	0.2496	0.6849	0.1129	**0.0060**
Uncertainty	0.0730	**0.0055**	**0.0059**	**0.0459**	**0.0303**	**0.0009**	**0.0028**
Joint	**0.0007**	**0.0223**	**0.0364**	0.2676	0.1926	$< \mathbf{10^{-4}}$	**0.0114**
5-year ahead Beliefs							
Regression	N.A.	N.A.	N.A.	N.A.	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	0.2981
Mean	N.A.	N.A.	N.A.	N.A.	**0.0134**	0.2218	0.7128
Variance	N.A.	N.A.	N.A.	N.A.	0.8124	0.2787	**0.0198**
Uncertainty	N.A.	N.A.	N.A.	N.A.	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	**0.0002**
Joint	N.A.	N.A.	N.A.	N.A.	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$

- The joint test rejects whenever the Regression Test rejects.
- One additional rejection for 5-year ahead beliefs in 2013.

Why Do We Find an Additional Rejection?

- Mean-beliefs are close;

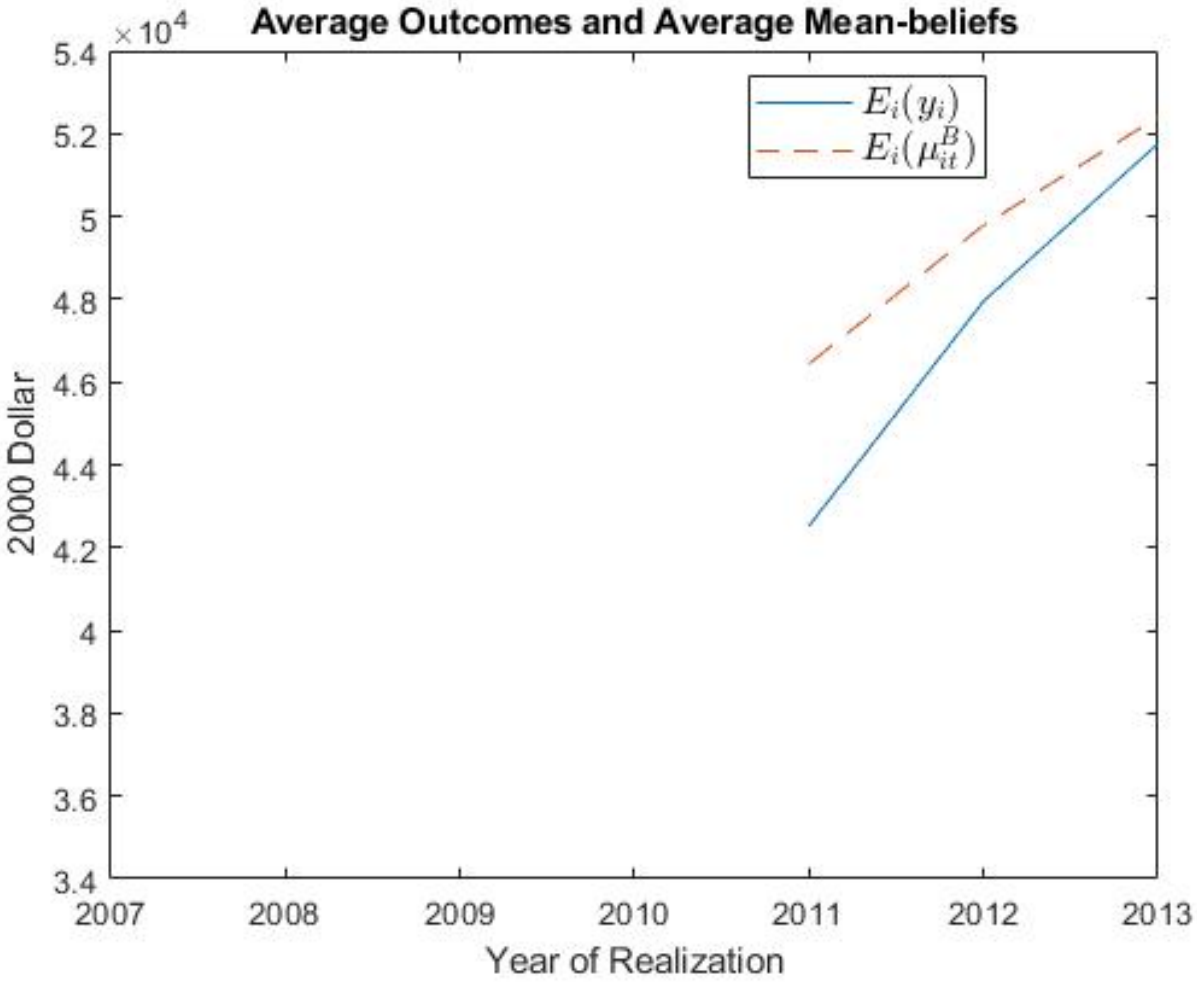

Figure 1: Descriptive Statistics (5-Year ahead Beliefs)

Why Do We Find an Additional Rejection?

- Mean-beliefs are close;
- Variances are vastly different.

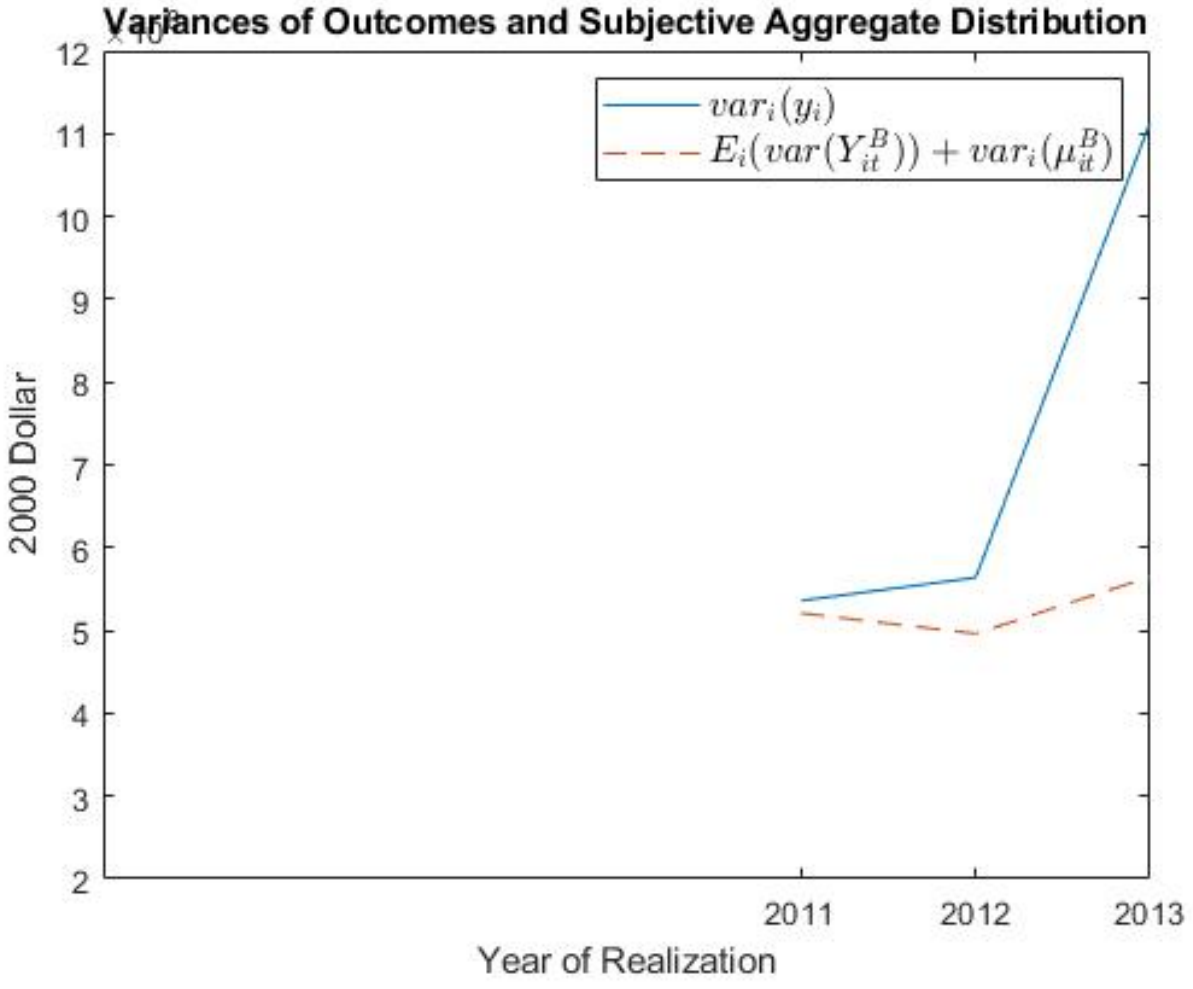

Figure 1: Descriptive Statistics (5-Year ahead Beliefs)

Why Do We Find an Additional Rejection?

- Mean-beliefs are close;
- Variances are vastly different.

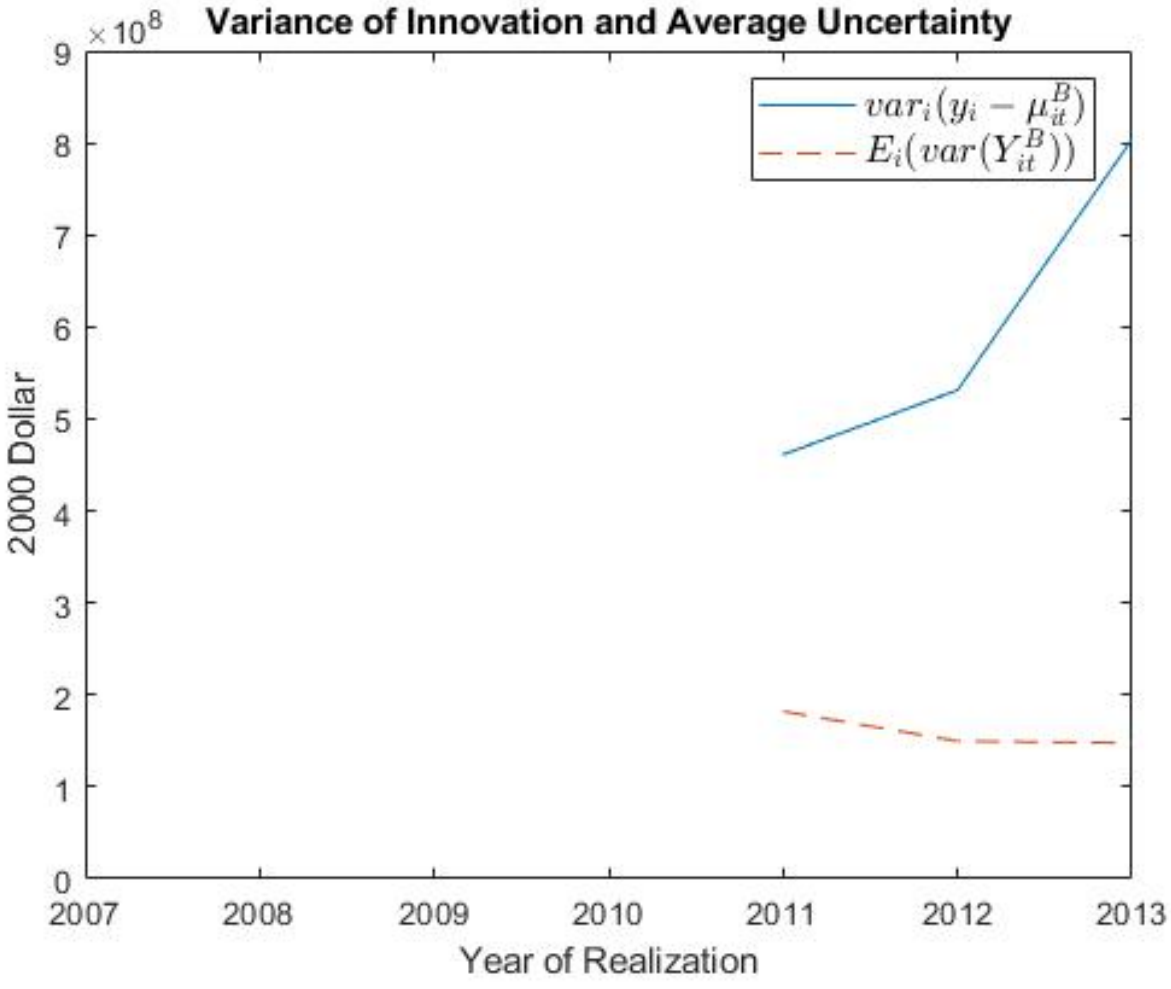

Figure 1: Descriptive Statistics (5-Year ahead Beliefs)

Testing in the Presence of Additive Aggregate Shocks

- We consider a case where Y_i depends on some additive aggregate shocks: $Y_i = Y_I(X_i) + Y_A(Z)$, $X_i \perp\!\!\!\perp Z$
  - commonly assumed in the (log-)income process literature (e.g., Meghir and Pistaferri, 2011).
- Previous tests are not valid as they are.
  - A particular realization of Z can arbitrarily shift the distribution of y_i.
  - Uncertainty about $Y_A(Z)$ is incorporated in belief distributions but not the distribution of y_i.

Testing in the Presence of Additive Aggregate Shocks

- We can adapt these tests to accommodate additive aggregate shocks.
  - based on Y_{it}^{B} is more dispersed than $Y_I(\mathrm{X}_i)|(\mathrm{X}_i^{\theta_i,t-} = \mathrm{x}_i^{\theta_i,t-}, \Theta_i = \theta_i)$

Moment-Based Aggregate-Distribution Test

$$E_i(v_{it}^{B}) + var_i(\mu_{it}^{B}) \geq var_i(y_i).$$

Uncertainty Test

$$E_i(v_{it}^{B}) \geq var_i(y_i - \mu_{it}^{B}).$$

Percentile-Based Joint-Distribution Test

$$E_i[I(y_i - E_i(y_i - \mu_{it}^{B}) \in (C_{it}^{r,B}, C_{it}^{r',B}))] \geq 0.01(r' - r), \text{ for any } r < 50 < r'.$$

- Our general empirical finding of many more rejections for in-school beliefs than for post-college beliefs remains. detail

Conclusion

- We develop several distributional tests of RE for beliefs elicited using probabilistic expectations questions.
- We show, both theoretically and empirically, our tests can detect more types of violations of RE than mean-based RE tests.
- Using our tests, we find post-college income beliefs are more "rational" than in-school income beliefs.
- Our empirical findings have implications for specifying income beliefs in structural models.

Thank you!

1-Year ahead Beliefs

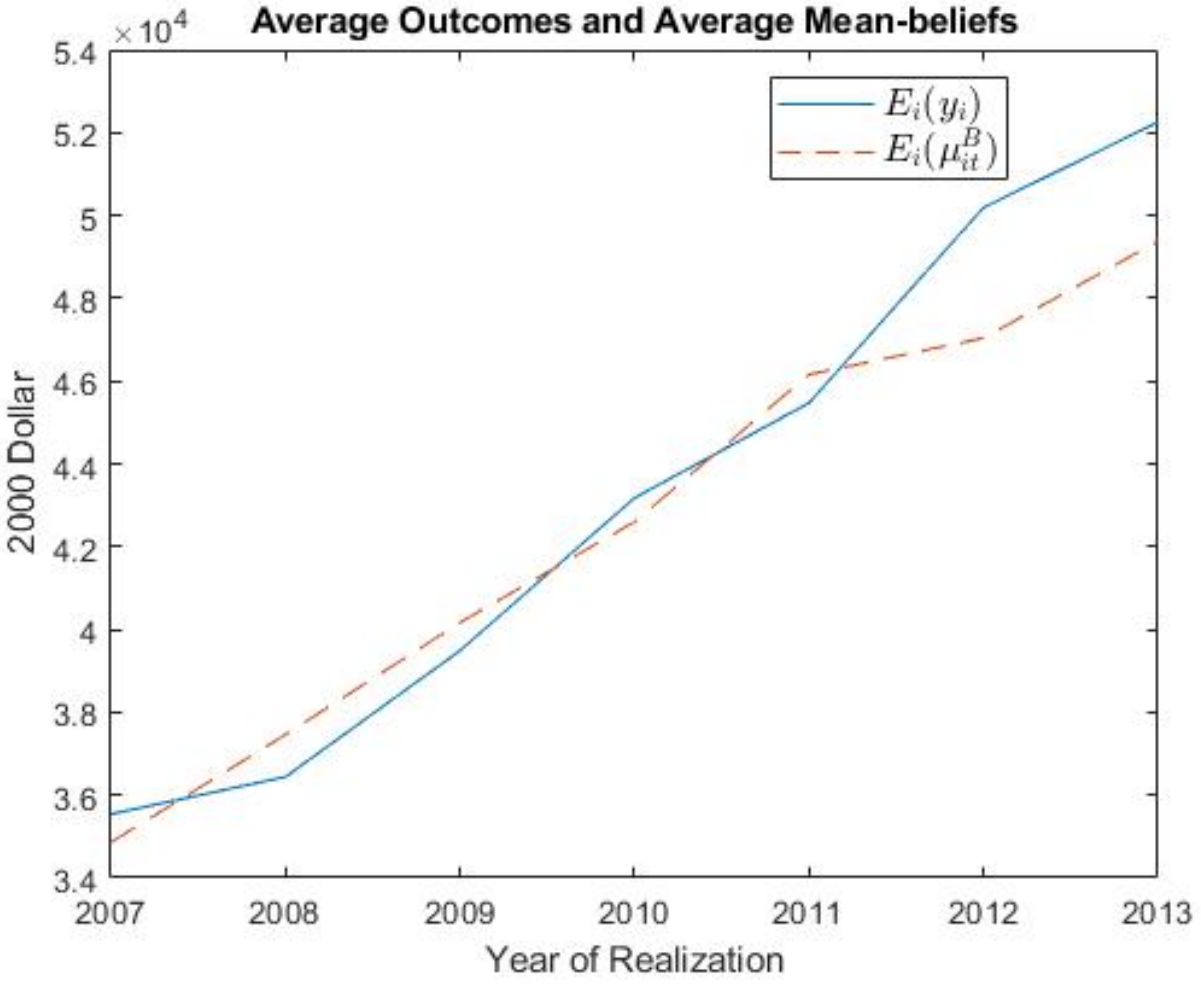

Figure 2: Descriptive Statistics (1-Year ahead Beliefs)

1-Year ahead Beliefs

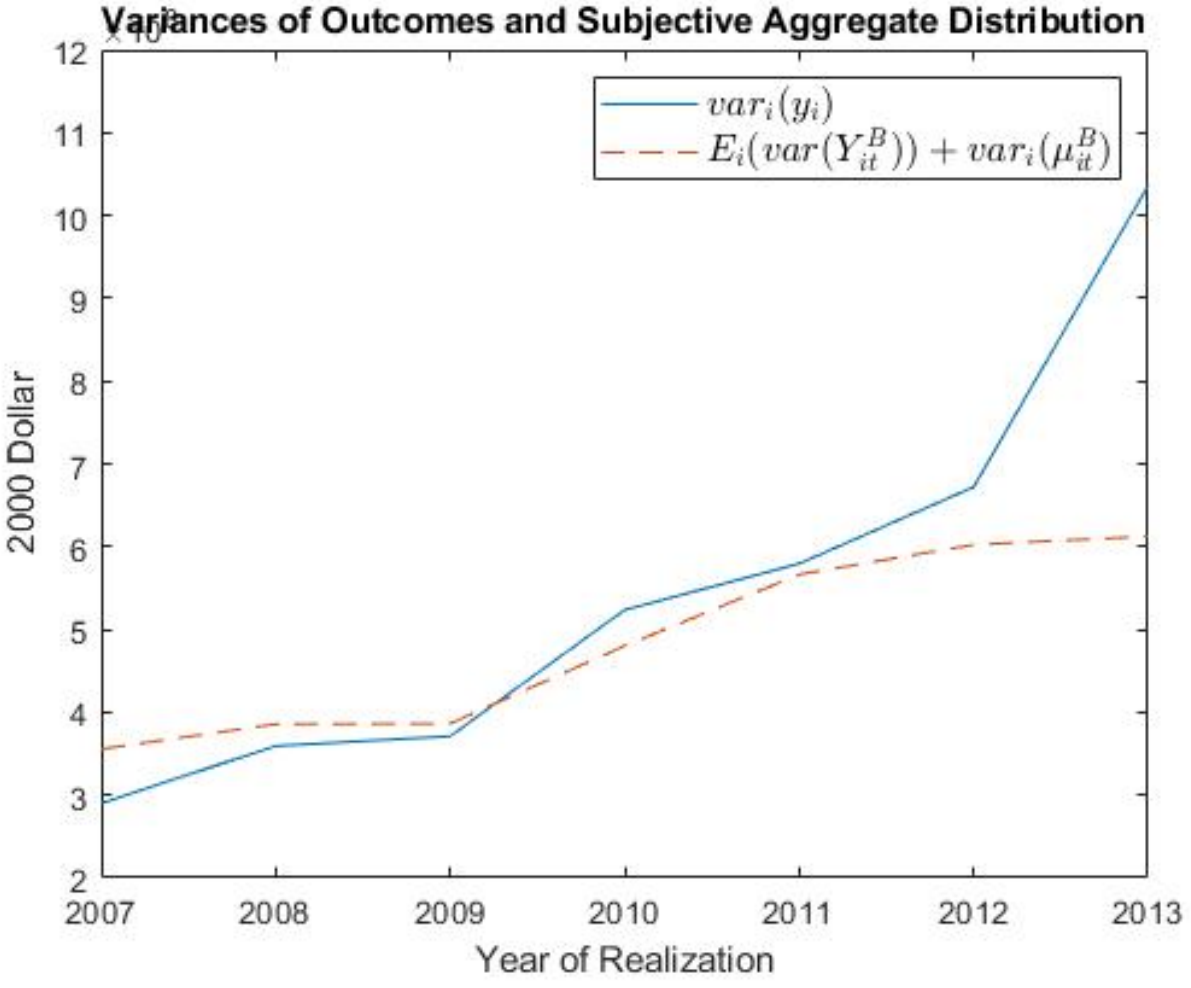

Figure 2: Descriptive Statistics (1-Year ahead Beliefs)

1-Year ahead Beliefs

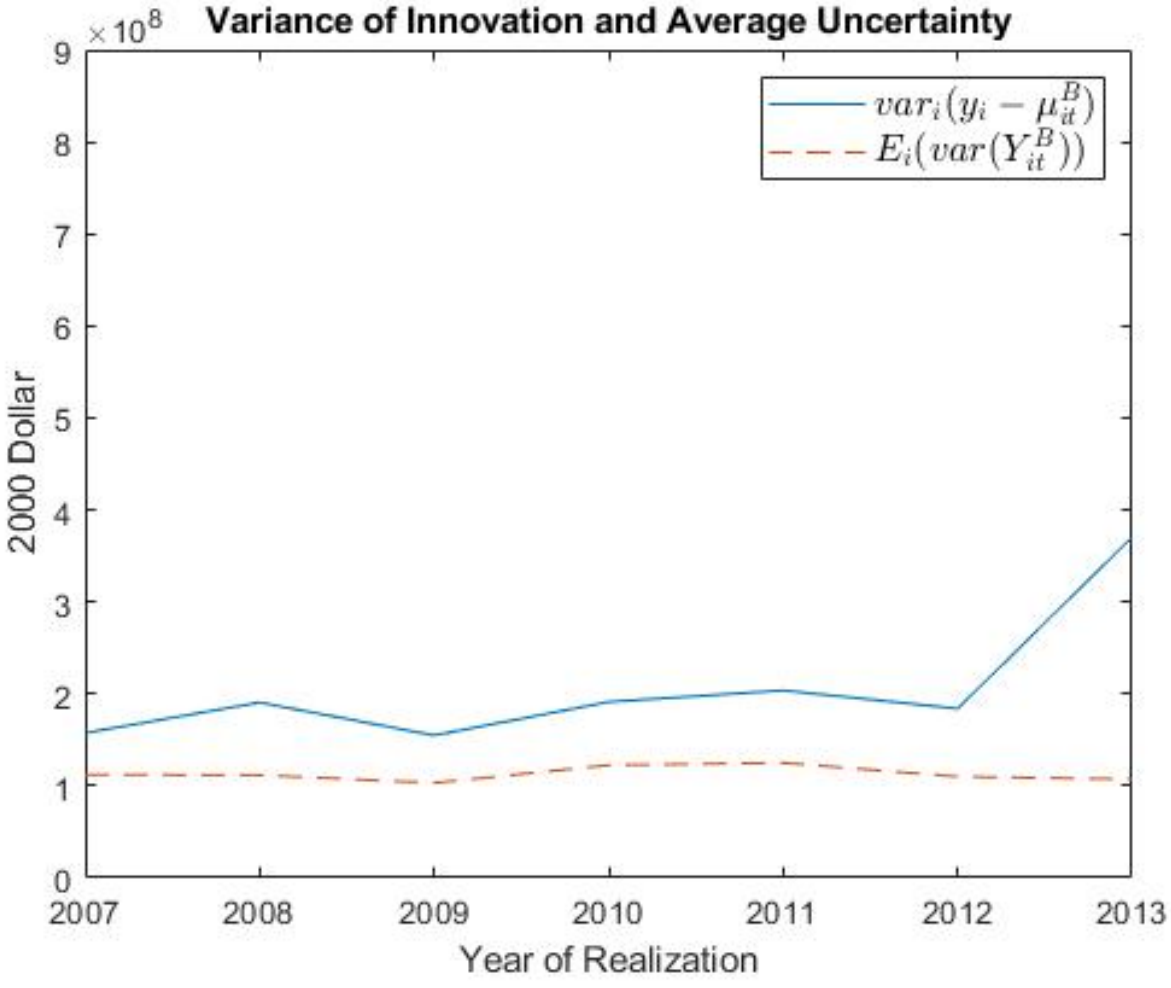

Figure 2: Descriptive Statistics (1-Year ahead Beliefs)

Testing Results in the Presence of Aggregate Shocks

Table 4: Rational Expectations Tests about Income at Age 28 with Aggregate Shocks (p-values)

10-year ahead	9-year ahead	8-year ahead	7-year ahead	6-year ahead	5-year ahead	1-year ahead
Aggregate-Distribution Test						
0.0978	**0.0138**	**0.0014**	**0.0125**	**0.0235**	0.9998	0.9940
Uncertainty Test						
$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	0.1461	0.8635
DGM Test						
1.0000	1.0000	1.0000	1.0000	1.0000	0.1543	0.5952
Regression Test: p-value						
$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	$< \mathbf{10^{-4}}$	**0.0093**

Notes: All tests are applied to balanced panels.

back

Income Expectations Questions - In-school

- We first ask you to indicate the lowest possible amount of money you might make and the highest amount of money you might make. We then ask you to divide the values between the lowest and the highest into four intervals. Please mark the intervals so that there is a 25% chance that your income will be in each of the intervals. Please write the **FIVE NUMBERS** that describe the income which you would expect to earn at the following ages or times under this hypothetical scenario.

I. Your income during the first full year after you leave school

lowest	highest

II. Your income at age 28 (note: if you are 20 years of age or older, give your income 10 years from now)

lowest	highest

III. Your income at age 38 (note: if you are 20 years of age or older, give your income 20 years from now)

lowest	highest

back

Income Expectations Questions - Post-college

- Think ahead to one year from today. Tell us the percent chance that the total yearly earnings for your family will be in each of the following categories **one year from today**. **NOTE: Each number should each be between 0 and 100 and the numbers should sum to 100.** You should enter a zero for a category if there is zero change your total family earnings will be in that interval next year.

Earnings Interval	**Percent Chance**
[$0, $15,000)	_____
[$15,000, $30,000)	_____
[$30,000, $45,000)	_____
[$45,000, $60,000)	_____
[$60,000, $75,000)	_____
[$75,000, $90,000)	_____
[$90,000, $105,000)	_____
[$105,000, $120,000)	_____
[$120,000, $135,000)	_____
[$135,000, 1 million)	_____

back