

Crossley, Thomas F.; Gong, Yifan; Stinebrickner, Todd R.; Stinebrickner, Ralph

Working Paper

Examining income expectations in the college and early post-college periods: New distributional tests of rational expectations

IFS Working Paper, No. W21/01

Provided in Cooperation with:

Institute for Fiscal Studies (IFS), London

Suggested Citation: Crossley, Thomas F.; Gong, Yifan; Stinebrickner, Todd R.; Stinebrickner, Ralph (2021) : Examining income expectations in the college and early post-college periods: New distributional tests of rational expectations, IFS Working Paper, No. W21/01, Institute for Fiscal Studies (IFS), London,
<https://doi.org/10.1920/wp.ifs.2021.0121>

This Version is available at:

<https://hdl.handle.net/10419/242900>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Thomas F. Crossley
Yifan Gong
Todd Stinebrickner
Ralph Stinebrickner

21/01
Working paper

Examining income expectations in the college and early post-college periods: new distributional tests of rational expectations



SPENCER

THE
ANDREW W.
MELLON
FOUNDATION

SSHRC  CRSH

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

Thomas F. Crossley* Yifan Gong[†] Todd Stinebrickner[‡]
Ralph Stinebrickner[§]

January 11, 2021

Abstract

Unique longitudinal probabilistic expectations data from the Berea Panel Study, which cover both college and early post-college periods, are used to examine young adults' beliefs about their future incomes. We introduce a new measure of the ex post accuracy of beliefs, and two new approaches to testing whether, ex ante, agents exhibit Rational Expectations. We show that taking into account the additional information about higher moments of individual belief distributions contained in probabilistic expectations data is important for detecting types of violations of Rational Expectations that are not detectable by existing mean-based tests. Beliefs about future income are found to become more accurate as students progress through school and then enter the post-college period. Tests of Rational Expectations almost always reject for the in-school period, but the evidence against Rational Expectations is much weaker in the post-college period.

Keywords: Subjective Expectations, Rational Expectations, Test, Income Uncertainty
JEL codes: D84

*Department of Economics, European University Institute, Fiesole, FI, 50014, Italy and Institute for Fiscal Studies, London. (email: thomas.crossley@EUI.eu).

[†]Department of Economics, University of Western Ontario, London, ON N6A 3K7, Canada (email: ygong48@uwo.ca).

[‡]University of Western Ontario, London, ON N6A 3K7, Canada (email: trstineb@uwo.ca).

[§]Department of Mathematics, Berea College, Berea, KY 40403, United States (email: stinebricknerr@berea.edu).

1 Introduction

Whether young adults have reasonable expectations about their future incomes is important for a variety of policy reasons. For example, whether college attendance decisions are optimal depends on the quality of beliefs about the financial returns to college education. Gong, Stinebrickner, and Stinebrickner (2020) show that these decisions also depend on the second moment of beliefs (uncertainty). Beyond schooling, whether consumption and savings decisions are optimal may depend on the quality of beliefs about future income (Sandmo, 1970; Leland, 1978; D’Haultfoeuille, Gaillac, and Maurel, 2018).

This paper takes advantage of data from the Berea Panel Study (BPS) to study young adults’ beliefs about their future incomes. As described in Section 2, the BPS follows two cohorts of students from college entrance to about age 30. The BPS was one of the first longitudinal studies to use probabilistic expectations questions to elicit, for each respondent, information about the full distribution describing beliefs (Stinebrickner and Stinebrickner, 2012). In each year, the study elicited each student’s beliefs about post-college income at particular future points in time. Subsequently, the Study then also collected each student’s actual realized earnings at those same points of time. The BPS is unique in allowing a comparison of probabilistic expectations and outcomes for an extended period, from the time of college entrance through the early part of the working life.

We begin, in Section 3, by characterizing the *ex post* accuracy of respondents’ beliefs, which naturally involves measuring some type of average distance between a person’s subjective beliefs about an outcome and the realization of the outcome. We develop a novel average belief-weighted mean squared error measure of accuracy and show how accuracy evolved through college and into the early working lives for these young adults. Our measure directly exploits the BPS information about belief distributions. It decomposes naturally into a component that indicates the average mean bias and two components that indicate whether respondents had, on average, an appropriate degree of uncertainty. Our *ex post* accuracy measure suggests that beliefs become more accurate as students progress through school and then enter the post-college period. *Ex post*, respondents were too optimistic (positively biased) while in college, and they were more uncertain while in college than afterwards.

We turn to formally testing whether, *ex ante*, beliefs satisfy Rational Expectations in Sections 4 through 6. The Rational Expectations Hypothesis requires that the distribution describing a person’s subjective beliefs about future income is equal to the actual conditional

distribution of future income given the person’s current information set. It is worth noting that, while the concepts of ex post accuracy and ex ante rationality are closely related, they convey different information about the quality of income beliefs. For example, if people tend to have rational but inaccurate beliefs about future income, then policy may seek to reveal additional information about factors such as ability that create uncertainty when they are not completely observed. On the other hand, if beliefs tend to be irrational, then policy might seek to help agents use the information they have more effectively, for example by providing guidance about how known factors such as educational attainment or college major influence the distribution of future income.

Existing tests of Rational Expectations typically use only the mean of a respondent’s subjective belief distribution (see Pesaran and Weale, 2006, for a survey). For example, researchers commonly use a Regression test, which focuses on the coefficients in a regression of outcomes on mean beliefs. As such, it does not require information on higher moments of the belief distribution. We make an important methodological contribution by recognizing a central implication of Rational Expectations - that agents’ beliefs are not only unbiased, but also, for example, exhibit an appropriate degree of uncertainty. In Section 4, we develop two new approaches for testing Rational Expectations on micro data when probabilistic expectations about a continuous variable are available. Unlike many traditional tests that are based entirely on mean beliefs, our proposed tests also assess the second and higher moments of agents’ beliefs.

Probabilistic expectations data, which have become increasingly common, can be collected in two ways: by eliciting percentiles of the belief distribution or by eliciting the perceived probabilities for some partition of the realization space (Manski, 2018). We refer to the former as the *percentile* format and the latter as the *probability* format. The BPS collects expectations data in both ways. Our first new approach to testing Rational Expectations develops naturally from our measure of ex post accuracy, testing whether beliefs are, on average, unbiased, and whether subjective uncertainty captures the objective variance of innovations. We refer to this test as the “Mean-Variance” Test. Like many existing tests of Rational Expectations, this test requires panel data.

Our second new testing strategy takes a quite different, and to the best of our knowledge, entirely novel approach. It builds on the insight that, under Rational Expectations, the distribution that results from aggregating the subjective beliefs of individuals will be identical to the distribution of realizations (in the absence of aggregate shocks). We refer to tests based

on this approach as “Aggregate-Distribution” tests. In addition to assessing the higher moments of agents’ beliefs, this approach has several further advantages: 1) it results in *repeated cross-section* tests, so it does not require panel data; 2) when belief data is elicited in the probability format, a fully nonparametric test can be deployed; 3) this approach is robust to some kinds of reporting error in survey data on beliefs, unlike, for example, the Regression test, which is quite sensitive to measurement error in mean beliefs.

Intuitively, taking into account the additional information contained in the higher moments of individual belief distributions should allow us to detect more types of violations of Rational Expectations. An important part of our methodological contribution comes from demonstrating that this is true for our new tests. As one example, we show that the null hypotheses of the Mean-Variance Test and the Aggregate-Distribution Test jointly imply the null of the Regression Test, but that the reverse is not true.

In Section 5, we apply our two new approaches, as well as mean-based tests from the literature, taking advantage of the unique features of the BPS. One primary substantive contribution comes from the finding that Rational Expectations provides a better description of income beliefs for the post-college period than for the in-school period. Across different tests and different years of the in-school period, we almost always reject the null that individuals have Rational Expectations about income at age 28. However, our conclusions are substantially different in the post-college period, where we frequently do not reject the null of Rational Expectations about income at the age of 28. The BPS data also allow us to provide evidence about the practical importance of our theoretical results about the relationship between our new tests, which take into account higher moments, and existing mean-based tests. Consistent with these theoretical findings, we find that our tests reject more often than existing mean-based tests in the post-college period.

As with most existing tests of Rational Expectations, our proposed tests compare elicited beliefs to realizations from a single period. It is well known that Rational Expectations tests of this type are only strictly valid in the absence of aggregate shocks (see, e.g., Attanasio and Augsburg, 2016). In Section 6, we extend our analysis by considering how different tests (both ours and tests from the previous literature) of Rational Expectations can accommodate the possibility of additively separable aggregate shocks. This form of aggregate shocks is commonly assumed in the income process literature (see, e.g., Meghir and Pistaferri, 2011). We then apply modified versions of the tests to the BPS data. When we allow for additively separable aggregate shocks we continue to reject the rationality of in-school

beliefs for almost any combination of test and year of belief elicitation. These results stand in stark contrast to the results from the post-college period, where we rarely reject the null of Rational Expectations. Thus, importantly, the general pattern of many more rejections of Rational Expectations during college than after college is robust to allowing for aggregate shocks.

Related Literature: Our analysis of students' and young adults' beliefs about future income adds to a growing literature on students' beliefs and expectations (for example Zafar, 2011; Arcidiacono, Hotz, and Kang, 2012; Stinebrickner and Stinebrickner, 2012), and more generally to the broader literature on the use of probabilistic expectations data to study economic behaviour (see Manski, 2004; Delavande, 2008; Van der Klaauw, 2012; Manski, 2018).

Dominitz and Manski (1996), Dominitz and Manski (1997b) and Dominitz and Manski (1997a) are important early papers studying survey data on probabilistic expectations about income; Dominitz and Manski (1996) in particular focuses on students. The main focus in those papers is to describe the respondents' belief distributions. Guiso, Jappelli, and Terlizzese (1992) and Das and Donkers (1999) use similar data, also to characterize income uncertainty. Dominitz (1998) employs panel data with both probabilistic expectations and realizations of income data. He examines several implications of the Rational Expectations Hypothesis for such data, but does not provide formal tests. He also assumes throughout that realizations are independent, meaning that there is no possibility of aggregate shocks. Das and Van Soest (1997) and Das, Dominitz, and Van Soest (1999) conduct similar exercises, but with qualitative data on income expectations, rather than elicited probabilistic expectations.

A paper that is especially relevant is D'Haultfoeuille, Gaillac, and Maurel (2018), which also proposes a new test of the Rational Expectations Hypothesis using micro data on income. That paper quantifies the consequences of deviations from Rational Expectations in the context of a life-cycle model of consumption, providing a general motivation for the focus on the testing of Rational Expectations in our paper. However, perhaps of more specific relevance is that our testing approaches complement their test. Traditionally, it has been common for surveys to elicit only point expectations from respondents and researchers often interpret an elicited point expectation as the mean of a respondent's belief distribution. Adopting this interpretation, D'Haultfoeuille, Gaillac, and Maurel (2018) develop a repeated cross-section mean-based test that takes into account the implications of Rational Expectations on all (cross-sectional) moments of mean-beliefs and outcomes. However, in-

creasingly researchers and survey designers have turned to eliciting more information about belief distributions, particularly with a goal of understanding the subjective uncertainty that economic agents face. Our new panel test and our new repeated cross-section tests deal explicitly with this case. An important point highlighted by both the testing approaches we develop is that elicited probabilistic expectations provide information about the rationality of beliefs beyond that which is contained in point expectations. This is true even when agents' information sets are not observed. In our conclusion (Section 7), we also discuss the implications of our findings for the relevance of our findings for the design of survey questions that elicit probabilistic expectations.

2 Data

2.1 Berea Panel Study

Designed and administered by Todd Stinebrickner and Ralph Stinebrickner, the Berea Panel Study (BPS) is a multipurpose longitudinal survey project, which began in 1999 and collected detailed information for understanding a wide variety of issues related to higher education and the early part of the post-college period. The BPS has been used to study dropout, major choice, tasks, etc. (Stinebrickner and Stinebrickner, 2012, 2014a,b; Stinebrickner, Stinebrickner, and Sullivan, 2019).

The BPS took place at Berea College, which is located in central Kentucky. Berea operates under a standard liberal arts curriculum and students at Berea are similar in academic quality, for example, to students at the University of Kentucky (Stinebrickner and Stinebrickner, 2008).¹ Further, academic decisions and outcomes at Berea are similar to those found elsewhere. Stinebrickner and Stinebrickner (2003) shows that dropout rates are similar to the dropout rates at other schools (for students from similar backgrounds) and Stinebrickner and Stinebrickner (2014b) shows that patterns of major choice and major-switching are similar to those found in the NLSY by Arcidiacono (2004).

However, Berea has some features that are advantageous for research purposes. Perhaps most prominently, Berea operates under an objective of providing educational opportunities to “students of great promise, but limited economics resources.” Generally, the low income demographic group covered by the BPS is of particular interest to policymakers because, for

¹The average ACT score of students at Berea College is 23. In terms of demographic characteristics, approximately 60% of students are female and approximately 18% of students are black.

example, they are much less likely to graduate from college than students from more privileged backgrounds (Manski, 1992; Bowen and Bok, 1998; Bowen, Chingos, and McPherson, 2009). Further, issues related to the accuracy and rationality of beliefs studied in this paper may be of particular importance for this group. Students at Berea may be especially likely to suffer from information frictions since less than half (34%) of students have a parent who completed a college degree and, on average, the income of their families is roughly \$26,000 at the time of college entrance (Stinebrickner and Stinebrickner, 2014a). As always, it is important to keep in mind that caution is warranted when thinking about exactly how our results might generalize to, for example, students from higher income backgrounds.

The BPS consists of two cohorts. Baseline surveys were administered to the first cohort immediately before it began its freshman year in the fall of 2000 and baseline surveys were administered to the second cohort immediately before it began its freshman year in the fall of 2001. In total, 720 students answered this initial survey. The BPS was unique in its frequency of contact with respondents. Students who completed the baseline were eligible to participate in the 10-12 surveys that took place each year during the in-school period. We focus on the baseline survey and the surveys that were administered at the end of each academic year because these surveys contain the questions about income beliefs that we use. All students in the two entering cohorts who graduated were eligible for the annual post-college surveys that were administered in May/June until 2013, providing up to 9 years of post-college data.

2.2 Survey Questions

In-school Income Beliefs

The survey questions eliciting beliefs about future income while respondents were in school are of the percentile format. Specifically, as seen in Question A1 in Appendix A, the in-school survey questions elicited the minimum, the maximum, and the three quartiles of the distribution describing a student's beliefs about own annual income at three different ages (first year after graduation, age 28, and age 38).² In this paper we focus on beliefs about own income at age 28 because this is the age for which we have the most complete outcome data; respondents were younger than age 38 when the post-college survey ended, and a substantial number of students had not graduated in time to receive a post-college survey at the age of 23.

²We exclude observations for which the maximum is higher than \$250,000.

Detailed instructions were included with the survey. The first portion of the instructions involved a discussion of an example that would seemingly be of particular salience to students - beliefs about grade performance. This portion of the instructions began by noting that students might not be certain about future outcomes of interest. It then focused on helping students understand the concept of quartiles by, for example, describing a hypothetical set of responses in which the first and third quartiles of a student’s belief distribution (about grade performance on a hypothetical test) are not necessarily the same distance from the median. The second portion of the instructions described information specifically relevant for providing information about income beliefs. Of particular note, the survey instructions highlighted one fundamental reason that uncertainty might exist about future income by instructing students to take into account the possibility that they would be working part-time, working full-time or not working at all at a particular future age.

Given data collected in the percentile format, a distributional assumption is required to calculate moments of belief distributions. We employ a step-wise uniform distribution, though the results reported below are robust to alternative choices. Details of these calculations are provided in Appendix B. It is worth noting that when point expectations are collected in surveys, researchers often interpret responses as the mean (mathematical expectation) of the respondent’s belief distribution, but it is not clear that this is how the question was intended by survey designers or interpreted by respondents.³ They might for example, report the median or the mode (D’Haultfoeuille, Gaillac, and Maurel, 2018; Benítez-Silva et al., 2008). Thus, an assumption is also required to obtain a moment, such as the mean of the belief distribution, from point expectations questions.

We exclude observations for which the mean of the belief distribution is more than three sample standard deviations away from the cross-sectional sample mean of this object. Similarly we exclude observations for which the standard deviation of the belief distribution is more than three sample standard deviations away from the cross-sectional sample mean of this object.

Table 1 reports some descriptive statistics for students’ in-school belief distributions. As shown in the first three rows of the table, on average students downward adjust the minimum, the median, and also the maximum of their belief distributions monotonically over time. The fourth row of Table 1 shows that the interquartile range, which is a measure of uncertainty, is also decreasing over time.

³For example, the Federal Reserve Bank of New York’s Survey of Consumer Expectations has asked

Post-College Income Beliefs

The survey questions eliciting post-college beliefs about future income differ in that they ask about family income and in that they use the probability format.⁴ Specifically, as seen in Survey Question A2 in Appendix A, the post-college survey questions elicited the perceived probability that a respondent’s annual family income in the next year would fall in each of ten income categories and the perceived probability that a respondent’s annual family income in five years would fall in each of ten income categories. As with the in-school expectations questions, respondents were instructed to take into account the possibility of changes in job situations. Given the family nature of the survey question, respondents were also asked to take into account the possibility that the number of workers in the family might change due to, for example, marriage and the birth of children.

A distributional assumption is again required to derive moments of each respondent’s belief distribution from their responses. As detailed in Appendix B, we again employ a step-wise uniform distribution within bins, treating the maximum value in the top bin as 150,000 dollars. As before, we exclude observations for which the mean (or standard deviation) of the belief distribution is more than three sample deviations away from the cross-sectional sample mean of this object.

We find that there are relatively few respondents who assign 100 percent probability to one income bin and zero probabilities to other income bins. This percentage of respondents ranges between 5% and 8% during the post-college years. Looking across all post-college years, on average, respondents reported positive probabilities for about four income bins.

Due to space considerations, we do not show the full descriptive statistics for all reported probabilities. Instead, we note that, over time, individuals update, in a positive direction, their beliefs about income in the next year and in five years. For example, over time, the percentage of respondents who reported a positive probability for the bottom bin decreases, from 31% in 2006 to 20% in 2013. Similarly, over time, the percentage of respondents who reported a positive probability for the top bin increases, from 8% in 2006 to 23% in 2013.

Realizations

“What do you believe your annual earnings will be in four months?”

⁴The choice between the probability format and the percentile format is in part a trade off between the possibility of anchoring with the probability format (Dominitz and Manski, 1996, 1997b; Manski, 2004) against the likelihood that the probability format is easier for subjects to comprehend. Other expectations questions in the in-school survey (notably about grades) employed the probability format. The decision to move to the probability format in the post-college period was motivated, in part, by feedback from participants and in-school analyses suggesting that students were especially comfortable with this format.

Annual own and spousal income are constructed from Question A3 in Appendix A, which gave respondents flexibility over whether earnings were reported for an hourly, weekly, monthly, or yearly period. Annual family income is then constructed by adding up own and spousal income. All beliefs and realizations are deflated by the CPI to 2001 dollars. We exclude observations for which realized income is more than three sample deviation away from the cross-sectional sample mean of realized income. Descriptive statistics for income realizations are of direct relevance to the tests that we propose. Hence, we show them along with the empirical results of our tests in Section 5.

2.3 Data Quality

The baseline survey, which took place immediately after students arrived for their freshman year, was completed in paper format in the presence of Todd and Ralph Stinebrickner. Before completing the survey, students received extensive classroom training, which focused on delivering and explaining the instructions for the income expectations questions. Subsequent in-school surveys were distributed through the campus mail system and returned to Ralph Stinebrickner in person. The in-school approach to surveying was motivated partly by the potential benefit to response rates of providing students with compensation immediately after they completed surveys. However, this approach also allowed students to provide immediate feedback about survey comprehension and to receive further encouragement to complete the surveys in a conscientious manner.

The spirit of this rapid review and compensation approach was carried forward into the post-college survey period. Compensation was included in advance with the mailed surveys. Surveys were reviewed as soon as they were returned, and interaction with respondents followed immediately if necessary. Students had prior experience with the probability format used in the post-college income expectations questions, as it was used, for example, to elicit beliefs about grade performance in the in-school surveys.

This survey approach led to high response rates. Approximately 85% of all students who entered Berea in 2000 and 2001 completed the baseline surveys. Eligibility for subsequent in-school surveys depended on completing the baseline survey and remaining in college, and response rates for those surveys remained around 85%. In part because of the long-term relationship built with respondents and in part because substantial effort was made to build a comprehensive database of addresses, response rates remained high in the post-college period. All individuals belonging to the initial cohorts (whether or not they completed the

baseline survey) who graduated were eligible for the post-colleges surveys. Against this base, the response rates for the post-college survey remained between 80% and 90% for most of the post-college period.

This survey approach also produced data with very little item non-response and high internal consistency. As an example of the low item non-response, only 9 of the 722 respondents did not complete the income expectations question on the baseline in-school survey. As an example of the high internal inconsistency, we find that the percentage of respondents who do not report probabilities that add up to 100 in their responses to post-college income expectations questions is very close to zero in all post-college years (ranging from 2% to 5%).

More generally, the quality of expectations data in the BPS has been documented in previous work. As one example, Stinebrickner and Stinebrickner (2012) finds that a simple theoretical implication related to college dropout - that the dropout decision should depend on both a student's cumulative GPA and beliefs about future GPA - is satisfied when beliefs are directly elicited through survey questions, but is not satisfied when beliefs are constructed under a version of Rational Expectations. As a second example, Gong, Stinebrickner, and Stinebrickner (2019) propose a method for characterizing the amount of classical measurement error in responses to expectations questions, and find that a relatively small amount of measurement error is present in responses to in-school income expectations questions. Of course, as suggested by Manski and Molinari (2010), measurement error present in survey responses might not be classical if, for example, it arises because of rounding behavior. Consistent with much other research, we find some evidence of rounding. For example, pooling together students' responses to both the 1-year ahead and 5-year ahead income expectations questions for all post-college years, we find that, for 84% of the responses, the reported probabilities associated with all ten income bins are multiples of 5; for 49% of the responses, the reported probabilities are multiples of 10; for 20% of the responses, the reported probabilities are multiples of 25; and for 13% of the responses, the reported probabilities are multiples of 50. Another possibility, recently investigated by Giustinelli, Manski, and Molinari (2019) is that individuals may have imprecise beliefs. As the BPS, which began in 2000, did not have individuals report intervals for probabilities, we are unable to assess this directly. However, in our concluding remarks (Section 7), we discuss how belief imprecision could affect the interpretation of our findings, and possible extensions to deal with this possibility.

3 Descriptive Evidence on Ex Post Accuracy

3.1 A Measure of Ex Post Inaccuracy

Let Y_i denote the random variable associated with an outcome of interest. To make things concrete, in our context, Y_i will be either i 's own income or i 's family income at a particular future time. For the sake of clarity, when we describe our empirical results in later sections, we add subscripts O and F to all relevant objects to differentiate between own and family outcomes.

For each individual i in a sample, the realization of Y_i is denoted y_i . At an earlier time t , each individual i reports information of relevance for characterizing the distribution describing her beliefs about y_i . We let Y_{it}^B denote the random variable whose distribution describes the subjective beliefs of individual i at time t about y_i .

Our interest in this section is in examining the ex post accuracy of beliefs, in other words, some measure of how beliefs deviate from realized outcomes. For individual i , we propose the measure of ex post inaccuracy at time t , denoted Δ_{it}^B , given by:

$$\begin{aligned}\Delta_{it}^B &= E_{Y_{it}^B}[(Y_{it}^B - y_i)^2] \\ &= E_{Y_{it}^B}[(Y_{it}^B - \mu_{it}^B) - (\mu_{it}^B - y_i)]^2 \\ &= E_{Y_{it}^B}[(Y_{it}^B - \mu_{it}^B)^2 + (\mu_{it}^B - y_i)^2 - 2(Y_{it}^B - \mu_{it}^B)(\mu_{it}^B - y_i)] \\ &= \text{var}_{Y_{it}^B}[Y_{it}^B] + (\mu_{it}^B - y_i)^2,\end{aligned}\tag{1}$$

where $\mu_{it}^B = E_{Y_{it}^B}(Y_{it}^B)$. Δ_{it}^B is the ex post belief-weighted mean squared error in individual i 's beliefs. The last line of Equation (1) shows that this measure can be written as the sum of individual i 's uncertainty (measured by the individual variance of Y_{it}^B) and her squared ex post error (relative to her mean belief).

Aggregating Δ_{it}^B over i yields an aggregate measure of ex post inaccuracy of beliefs at time t , which we denote Δ_t^B :

$$\Delta_t^B = E_i\{E_{Y_{it}^B}[(Y_{it}^B - y_i)^2]\}.\tag{2}$$

Throughout the paper, we use an i subscript whenever we take the expectation across individuals, and use the corresponding individual-specific random variable as the subscript whenever we take the expectation over the distribution of that random variable for an indi-

vidual.

A better understanding of Δ_t^B comes from inserting Equation (1) into Equation (2):

$$\begin{aligned}\Delta_t^B &= E_i\{E_{Y_{it}^B}[(Y_{it}^B - y_i)^2]\} \\ &= E_i[\text{var}_{Y_{it}^B}[Y_{it}^B]] + E_i[(\mu_{it}^B - y_i)^2].\end{aligned}\tag{3}$$

In the usual mean squared error way, the second term can be written as the sum of a *cross-sectional* variance and a squared aggregate error:

$$\Delta_t^B = \underbrace{E_i[\text{var}(Y_{it}^B)]}_{\delta_{t,1}^B} + \underbrace{\text{var}_i(y_i - \mu_{it}^B)}_{\delta_{t,2}^B} + \underbrace{(\bar{y} - \bar{\mu}_t)^2}_{\delta_{t,3}^B},\tag{4}$$

where $\bar{\mu}_t = E_i(\mu_{it}^B)$, and $\bar{y} = E_i(y_i)$.

Equation (3) shows that our measure of ex post inaccuracy, Δ_t^B , consists of three components, each of which is of interest. The first component $\delta_{t,1}^B \equiv E_i[\text{var}(Y_{it}^B)]$ measures the average amount of subjective uncertainty about y_i at time t . The second component $\delta_{t,2}^B \equiv \text{var}_i(y_i - \mu_{it}^B)$ is the cross-sectional variance of innovations $(y_i - \mu_{it}^B)$ and measures the dispersion of relevant new information received after t . If individuals' realizations are i.i.d., this is the individuals' "actual" uncertainty about y_i at time t . The third component, the squared aggregate error, $\delta_{t,3}^B \equiv (\bar{y} - \bar{\mu}_t)^2$, captures the size of the systematic ex post prediction error.

It is worth noting in advance that the three individual components $\delta_{t,1}^B$, $\delta_{t,2}^B$, and $\delta_{t,3}^B$ are useful for reasons beyond their role in the computation and understanding of Δ_t^B . As we discuss in detail in Section 4.2, the three separate components provide the foundation for one of our new tests of the Rational Expectations Hypothesis. Hence, this decomposition of Δ_t^B builds a link between the ex post accuracy and ex ante rationality of subjective beliefs.

3.2 Results

We begin by examining the ex post accuracy of income beliefs during college. We take advantage of the in-school survey questions described in Section 2.2 that elicit the minimum, maximum, and the three quartiles of the subjective belief distribution about own income at age 28, $Y_{O,it}^B$. As detailed in Appendix B, fitting the elicited data to a step-wise uniform distribution yields values of $E(Y_{O,it}^B)$ and $\text{var}(Y_{O,it}^B)$ for each student. These objects along with

realizations $y_{O,i}$ allow us to calculate the three components of ex post inaccuracy measure, $\delta_{O,t,1}^B$, $\delta_{O,t,2}^B$, and $\delta_{O,t,3}^B$. Then, the inaccuracy measure for own income at age 28, $\Delta_{O,t}^B$, can be obtained from Equation (3).

A unique feature of the BPS is that it elicited beliefs at the time of entrance and at the end of each year of college. For t ranging from 10 years before age 28 (at college entrance) to 6 years before age 28 (around the end of college), Panel A of Table 2 shows $\delta_{O,t,1}^B$, $\delta_{O,t,2}^B$, $\delta_{O,t,3}^B$, and $\Delta_{O,t}^B$, where the sample for a particular t consists of students for which both income expectations at time t and annual income at age 28 are observed.⁵

Consistent with the notion that students learn about income-influencing factors during college, Column 4 of Panel A in Table 2 shows that the ex post inaccuracy measure Δ_t decreases in a monotonic fashion during college. A further examination of the first three columns indicates that this is also true for all three components of Δ_t . Between college entrance (10-year ahead) and the end of college (6-year ahead), the measure of aggregate subjective uncertainty, $\delta_{O,t,1}^B$, drops by about 40%, from 123.6 to 72.3.⁶ Similarly, we find that the measure of aggregate actual uncertainty, $\delta_{O,t,2}^B$, decreases during college by roughly 60%, from 442.9 at college entrance to 178.3 at the end of college. However, most striking is the third column of Panel A, which shows that the systematic bias about $y_{O,i}$, as characterized by $\delta_{O,t,3}^B$, falls by about 90% during college. At college entrance, students overestimate $Y_{O,i}$ by $\sqrt{324.1} = 18.00$ thousand dollars. At the end of college, this overestimation is only $\sqrt{35} = 5.92$ thousand dollars.

Given the substantial changes in the inaccuracy measure that take place even during school, it is natural to consider also the path of inaccuracy after school, when direct involvement in the job market may provide students with significant additional information. Providing evidence about inaccuracy in the post-college period necessitates taking into account that BPS expectations questions were somewhat different in the schooling and post-schooling periods. Perhaps most notably, the post-college questions elicit beliefs about future family annual income rather than future own annual income. Thus, before characterizing inaccuracy in the post-college period, we set the stage for an informative comparison of in-school and post-college beliefs by considering how the in-school inaccuracy measures about future own income in Panel A can be extended to characterize in-school inaccuracy about future

⁵We choose not to restrict our sample to those who report income expectations at all t because the sample size would be very small (slightly more than 100 students) in this case. Nonetheless, we find qualitatively similar results when we restrict our analysis sample in this way.

⁶This is broadly consistent with the findings about income uncertainty in Gong, Stinebrickner, and Stinebrickner (2019).

Table 1: Average Responses to In-college Expectations Questions about Own Annual Income at Age 28

Time of Belief-elicitation	Beginning of Year 1	End of Year 1	End of Year 2	End of Year 3	End of College
Average Minimum	33.28 (15.71)	32.53 (16.25)	29.86 (12.32)	27.25 (11.23)	23.34 (10.21)
Average Median	47.25 (19.32)	46.95 (19.16)	42.25 (15.41)	39.55 (14.62)	34.69 (12.47)
Average Maximum	64.04 (28.91)	63.43 (27.59)	57.71 (23.78)	54.14 (23.49)	48.64 (19.92)
Avg. Interquartile Range	14.08 (10.66)	14.05 (10.15)	12.58 (8.91)	12.13 (9.62)	11.04 (7.31)
Sample Size	667	581	414	371	361

Notes: All numbers are in units of 1,000s of (2001) U.S. dollars. Standard deviations are reported in the parenthesis.

Table 2: Ex-post Accuracy: Beliefs in School and after College

Item	$\delta_{t,1}^B$	$\delta_{t,2}^B$	$\delta_{t,3}^B$	Δ_t	Sample Size
Panel A: In School (Own Income at 28/29)					
10-year ahead	123.6	442.9	324.1	890.6	271
9-year ahead	109.8	333.4	302.5	745.7	259
8-year ahead	101.4	310.8	169.8	582.0	206
7-year ahead	87.1	271.0	111.3	469.4	202
6-year ahead	72.3	178.3	35.0	285.6	207
Panel B: In School (Family Income at 28/29)					
10-year ahead	581.5	1,486.2	1,487.4	3,555.0	299
9-year ahead	495.4	1,208.8	1,357.7	3,061.8	284
8-year ahead	420.0	989.4	880.3	2,289.7	228
7-year ahead	380.9	986.7	631.8	1,999.4	219
6-year ahead	299.9	687.3	340.0	1,327.2	227
Panel C: After College (Family Income at 28/29)					
5-year ahead	162.3	432.1	8.5	602.9	159
1-year ahead	117.1	180.4	0.1	297.6	289

Notes: The unit of annual income is \$1,000 USD. Panel A reports the ex post inaccuracy measures of students' beliefs in college about own annual income at age 28. Panel B reports the ex post inaccuracy measures of students' beliefs in college about annual family income at age 28. Panel C reports the ex post inaccuracy measures of students' beliefs after college about annual family income at age 28 (or 29).

family income.

The in-school family income inaccuracy measure again comes from Equation (3), with all objects of relevance now being related to the family. Realized family income at age 28, $y_{F,i}$, is observed. Then, what is required is to characterize the mean and variance of the distribution describing beliefs about future family income, $Y_{F,it}^B$.

Student i 's annual family income at age 28, $y_{F,i}$, is given by:

$$y_{F,i} = y_{O,i} + m_i y_{S,i}. \quad (5)$$

where m_i is a binary indicator for whether a student is married at age 28 and $y_{S,i}$ represents spousal income (if present) at age 28. Our ability to characterize beliefs about $y_{O,it}^B$ were described during the computation of Panel A. Then, what is needed is to characterize beliefs about spousal income as captured by the second term on the right side of Equation (5). With respect to characterizing beliefs about the first element in this product, importantly, the BPS elicited each student's perceived probability at time t of being married at age 28, P_{it}^M . What remains is to characterize spousal beliefs about $y_{S,i}$, $Y_{S,it}^B$. In Appendix C, we explore the assumptions that allow us to characterize these beliefs, and therefore beliefs about future family income, $Y_{F,it}^B$.

Panel B of Table 2 shows $\delta_{F,t,1}^B$, $\delta_{F,t,2}^B$, $\delta_{F,t,3}^B$, and $\Delta_{F,t}^B$ for the same range of t used in Panel A. While all of the measures are considerably larger than their counterparts in Panel A due to the presence of additional uncertainty related to future marital status and future spousal income, as expected, the results in Panel B exhibit the same general pattern as the results in Panel A; all three components of inaccuracy measure are dropping monotonically during college.

Turning to examining inaccuracy in the post-college period, the needed mean and variance of post-college beliefs about future family income, $E(Y_{F,it}^B)$ and $var(Y_{F,it}^B)$, can be obtained directly since students directly report beliefs about future family income. In terms of specifics, students report beliefs about annual family income one year and five years in the future. Thus, 5-year ahead beliefs about family income at age 28 are available for a student who completes a survey at age 23 and 1-year ahead beliefs about family income at age 28 are available for a student who completes a survey at age 27. Unfortunately, as noted earlier, many students are older than 23 when they receive their first post-college survey. As a result, we compute our 5-year ahead post-college inaccuracy measure by comparing 5-year ahead

beliefs elicited at age 24 to realized family income at age 29.

The fourth column of Panel C shows that the downward trend of the ex post inaccuracy measures that we see for in-school beliefs continues through the post-college period. Beliefs about future family income become substantially more accurate as students make the transition from college to the labor market. The 5-year ahead measure is less than half of the 6-year ahead measure from the fourth column of Panel B. Beliefs after entering the labor market further improve as age becomes closer to 28, as shown by the 1-year ahead measure.

The decomposition of our ex post accuracy measure in the first three columns of Panel B and Panel C provides evidence about why the substantial improvement in accuracy occurs in the post-college period. Most notably, the third columns of Panel B and Panel C indicate that the substantial overestimation of future income that is present in college dissipates quickly after labor market entry. At the end of college, students overestimate their annual family income at age 28 by $\sqrt{340} = 18.44$ thousand dollars, on average. At age 24, this overestimation is reduced to $\sqrt{8.5} = 2.92$ thousand dollars on average. By age 27, the overestimation ($\sqrt{0.1} = 0.32$) has disappeared.

Of course, these improvements in ex post accuracy may result from multiple causes. The post-college beliefs are for a shorter time horizon. Nevertheless, the sharp improvement from the 6-year ahead measure to the 5-year ahead measure strongly suggests that the first experience in the post-college labor market has an important effect. The improvement within the post-college period, from the 5-year ahead measure to the 1-year ahead measure may reflect either further learning about post-college career prospects or that idiosyncratic incomes are likely easier to predict once more is known about the aggregate state of the economy.

4 Testing Rational Expectations: Theory

In this section we develop two new approaches to testing Rational Expectations about micro-level outcomes. One develops naturally from our measure of ex post accuracy. The other takes a quite different, and to the best of our knowledge, entirely novel approach. It builds on the insight that, under Rational Expectations, the distribution that results from aggregating the subjective beliefs of individuals will be identical to the distribution of realizations (in the absence of aggregate shocks). Finally, we analyze how our two tests relate to existing tests in the literature.

4.1 The Rational Expectations Hypothesis

4.1.1 Environment

As above we consider a situation in which we have data on realizations, y_i , of the random variable of interest, and some data on the random variable Y_{it}^B describing individuals' beliefs about y_i at an earlier time t . This information may come from a balanced panel, an unbalanced panel, or a repeated cross-section. In any case, we assume that all samples are randomly selected from the same population. For an unbalanced panel, this implies missing observations on beliefs or realizations are Missing Completely at Random (MCAR). Formally we assume:

- A1 The sample for which both Y_{it}^B and y_i are measured, the sample for which only Y_{it}^B is measured, and the sample for which only y_i is measured are all random samples from the same population.

As noted above, survey responses to point expectations questions are often interpreted as the mean of Y_{it}^B , which we will henceforth denote by μ_{it}^B . Probabilistic expectations questions can come in the percentile format, which elicits the p -th percentile (e.g., minimum, maximum, and quartiles) of Y_{it}^B for finitely many p , or the probability format, which elicits the probability that $Y_{it}^B \in S_k$ for $k = 1, \dots, K$, where $\mathcal{S} = \{S_1, \dots, S_k, \dots, S_K\}$ denotes a partition of the support of Y_{it}^B . Our two new tests take advantage of probabilistic expectations data, while most earlier tests are based on point expectations data.

We assume that Y_i is determined by many underlying random variables. We decompose these random variables into the ones that are independently distributed across individuals (individual-specific factors), \mathbf{X}_i , and the ones that are shared by all individuals (aggregate factors), \mathbf{Z} . We further denote the ones that are observed by individuals before t as \mathbf{X}_i^{t-} and \mathbf{Z}^{t-} , and the ones that are observed by individuals after t as \mathbf{X}_i^{t+} and \mathbf{Z}^{t+} . We use lower case versions of \mathbf{X}_i and \mathbf{Z} to denote their realizations.

4.1.2 Definition of Rational Expectations

Given the environment described above, we consider the version of the Rational Expectations given by:

Definition 1. Let $F_{Y_{it}^B}(y)$ denote the CDF of Y_{it}^B , the random variable describing individual i 's subjective beliefs about y_i at t . Individual i has Rational Expectations about y_i if and only if $F_{Y_{it}^B}(y) = F_{Y_i|(\mathbf{X}_i^{t-}=\mathbf{x}_i^{t-}, \mathbf{Z}^{t-}=\mathbf{z}^{t-})}(y)$ for $y \in \text{supp}(Y_i|(\mathbf{X}_i^{t-}=\mathbf{x}_i^{t-}, \mathbf{Z}^{t-}=\mathbf{z}^{t-}))$.

This definition of Rational Expectations requires individual i 's beliefs about a random variable Y_i to be identical to the actual distribution of Y_i conditional on her information set at t , $(\mathbf{X}_i^{t-}=\mathbf{x}_i^{t-}, \mathbf{Z}^{t-}=\mathbf{z}^{t-})$.

In this section we follow much of the previous literature on expectations of micro outcomes in assuming that there are no aggregate shocks observed by individuals after t , i.e., that \mathbf{Z}^{t+} is degenerate for all individuals. In Section 6 we explore the robustness of our results to this assumption.

In the case where \mathbf{Z}^{t+} is degenerate, we can express Y_i as a function of \mathbf{X}_i^{t-} , \mathbf{X}_i^{t+} , and \mathbf{Z}^{t-} : $Y_i = Y(\mathbf{X}_i^{t-}, \mathbf{X}_i^{t+}, \mathbf{Z}^{t-})$. Since \mathbf{Z}^{t-} is common to all individuals and is realized before beliefs are elicited, its presence does not affect the test of Rational Expectations. With regard to \mathbf{X}_i , ex ante individuals have identical distributions of \mathbf{X}_i , but realizations vary across people, and this produces the variation in the empirical distribution of realizations, y_i .

Turning to the variation in beliefs, this arises for two reasons. We allow the information set at time t to vary across individuals, i.e., \mathbf{X}_i^{t-} might contain different subsets of \mathbf{X}_i for different i . We let the random variable Θ_i denote the type of individual i , with each possible realization θ_i corresponding to one possible two-set partition of the full information set \mathbf{X}_i . Two individuals with the same θ_i observe the same set of random variables at t (denoted as $\mathbf{X}_i^{\theta_i, t-}$) and are uncertain about the remaining random variables (denoted as $\mathbf{X}_i^{\theta_i, t+}$). However, they might have different realizations of $\mathbf{X}_i^{\theta_i, t-}$. Thus, variation in beliefs (Y_{it}^B) is produced by (i) differences in what part of \mathbf{X}_i is observed by t (realizations of Θ_i) and (ii) differences in the realizations of the variables ($\mathbf{X}_i^{\theta_i, t-}$) for each type, θ_i . Throughout the paper, we maintain the assumption that Θ_i and \mathbf{X}_i are independent.⁷

To summarize, we impose the following assumptions throughout Section 4:

A2 \mathbf{Z}^{t+} is degenerate for all individuals.

⁷This assumption is needed because our definition of Rational Expectations (Definition 1) does not require individuals to correctly use information contained in Θ_i . In other words, it does not require individuals to adjust their beliefs about the distribution of unobserved factors based on the fact that these factors have not been observed. We would not need this independence assumption if we add this requirement to Definition 1.

This assumption is trivially satisfied if all individuals observe the same set of random variables at t (Θ_i is degenerate). One interpretation of this assumption is that θ_i is non-stochastic. As an example of what this assumption rules out, it would be violated if individuals know the state of their firm at t only if the firm is doing badly. Importantly, we note that this independence assumption is required for many existing tests of Rational Expectations, such as the Regression Test that we will discuss in detail later in Section 4.4.

A3 Θ_i and \mathbf{X}_i are independent.

4.2 Ex Post Inaccuracy and a Mean-Variance Test of Rational Expectations

As shown in Section 3.1, the aggregate measure of ex post inaccuracy can be written as the sum of the average subjective uncertainty $\delta_{t,1}^B = E_i[\text{var}(Y_{it}^B)]$, the aggregate “actual” uncertainty $\delta_{t,2}^B = \text{var}_i(y_i - \mu_{it}^B)$, and the squared systematic ex post prediction error $\delta_{t,3}^B = (\bar{y} - \bar{\mu}_t)^2$. Theorem 1 shows that a natural test of the RE assumption can be based on these three components.

Theorem 1. *Under assumptions A1 to A3, if all individuals have Rational Expectations about Y_i (Definition 1), then*

$$\delta_{t,1}^B = \delta_{t,2}^B \text{ and } \sqrt{\delta_{t,3}^B} = 0.$$

Proof. For ease of exposition, we prove the theorem for the case where Θ_i and \mathbf{Z}^{t-} are degenerate. We simplify notation by suppressing the θ_i superscript. In Appendix D we prove Theorem 1, allowing for nondegenerate Θ_i and \mathbf{Z}^{t-} .

Under the assumption that individual i has Rational Expectations about Y_i , her beliefs at t about y_i , Y_{it}^B , are given by:

$$Y_{it}^B = Y_i | (\mathbf{X}_i^{t-} = \mathbf{x}_i^{t-}). \quad (6)$$

This implies that:

$$\mu_{it}^B = E(Y_{it}^B) = E(Y_i | \mathbf{X}_i^{t-} = \mathbf{x}_i^{t-}), \quad (7)$$

$$\text{var}(Y_{it}^B) = \text{var}(Y_i | \mathbf{X}_i^{t-} = \mathbf{x}_i^{t-}). \quad (8)$$

The population distribution of Y_{it}^B is produced by the realizations of \mathbf{X}_i^{t-} . Hence, the first component, $\delta_{t,1}^B = E_i[\text{var}(Y_{it}^B)]$, is given by:

$$\delta_{t,1}^B = E[\text{var}(Y_i | \mathbf{X}_i^{t-})]. \quad (9)$$

Similarly, the population distribution of y_i is produced by the realizations of \mathbf{X}_i . Hence,

the second component, $\delta_{t,2}^B = \text{var}_i(y_i - \mu_{it}^B)$, is given by:

$$\begin{aligned}
\delta_{t,2}^B &= \text{var}[Y_i - E(Y_i|\mathbf{X}_i^{t-})] \\
&= \text{var}\{E[Y_i - E(Y_i|\mathbf{X}_i^{t-})|\mathbf{X}_i^{t-}]\} + E\{\text{var}[Y_i - E(Y_i|\mathbf{X}_i^{t-})|\mathbf{X}_i^{t-}]\} \\
&= \text{var}(0) + E[\text{var}(Y_i|\mathbf{X}_i^{t-})] \\
&= \delta_{t,1}^B,
\end{aligned} \tag{10}$$

where the second line follows from the law of iterative variance.

The second part of Theorem 1 directly follows from the law of iterative expectations:

$$\sqrt{\delta_{t,3}^B} = \bar{y} - \bar{\mu}_t = E(Y_i) - E(Y_i|\mathbf{X}_i^{t-}) = 0. \tag{11}$$

□

Note the role of the absence of the possibility of aggregate shocks. In the second part of Theorem 1 the absence of aggregate shocks allows us to interpret a systematic ex post prediction error as an ex ante bias. In the first part of Theorem 1 it implies that each individual's subjective uncertainty should equal her variance of idiosyncratic shocks, and the appropriate average of these can be measured by the cross-sectional variance of innovations. While generally these implications will not remain with the possibility of aggregate shocks after belief elicitation, a modified version of the second part of Theorem 1 can still hold if we impose some commonly-assumed structure on the aggregate shocks. We discuss this in Section 6.

In order to develop a statistical test, we begin by noting that the computation of $\delta_{t,2}^B$ and $\delta_{t,3}^B$ requires information about both Y_{it}^B and y_i for the same sample of individuals (a balanced panel). We let n denote the size of this sample. Further, we let $\hat{\delta}_{t,j}^B$ denote the sample analog of $\delta_{t,j}^B$ for $j = 1, 2, 3$, $\hat{\sigma}_{12}$ denote the sample standard deviation of $\text{var}(Y_{it}^B) - [(y_i - \mu_{it}^B) - \sqrt{\hat{\delta}_{t,3}^B}]^2$, and $\hat{\sigma}_3$ denote the sample standard deviation of $y_i - \mu_{it}^B$. Under the following assumption on the distribution of random variables, we are able to define and characterize the asymptotic distribution of two test-statistics using the central limit theorem.

A4 Both $y_i - \mu_{it}$ and $\text{var}(Y_{it}^B) - [(y_i - \mu_{it}^B) - (\bar{y} - \bar{\mu}_t^B)]^2$ have finite second moments.

Theorem 2. (*Mean-Variance Test*) Under assumptions A1 to A4, if all individuals have Rational Expectations about Y_i (Definition 1), then

$$\hat{u}_{12} \equiv \frac{\hat{\delta}_{t,1}^B - \hat{\delta}_{t,2}^B}{\hat{\sigma}_{12}/\sqrt{n}} \xrightarrow{d} N(0, 1),$$

$$\hat{u}_3 \equiv \frac{\sqrt{\hat{\delta}_{t,3}^B}}{\hat{\sigma}_3/\sqrt{n}} \xrightarrow{d} N(0, 1).$$

Theorem 2 suggests two different tests of the Rational Expectations Hypothesis, one based on the asymptotic two-tailed p-value of the test statistic \hat{u}_{12} and another based on the asymptotic two-tailed p-value of the test statistic \hat{u}_3 . Of course a joint test is possible, but having two tests may help differentiate between two different types of violations of the RE assumption. A rejection of the test based on \hat{u}_3 provides evidence of a systematic mean bias (mean predictions about y_i are not correct, on average). In contrast, a rejection of the test based on \hat{u}_{12} provides evidence that individuals do not have the “correct” amount of uncertainty about y_i .

4.3 Aggregate-Distribution Tests of Rational Expectations

The Mean-Variance Test developed in Section 4.2 requires that beliefs and outcomes be observed for the same set of individuals (a balanced panel). In this section, we propose Aggregate-Distribution Tests that can be implemented with unbalanced panels and repeated cross-sections. The Mean-Variance Test and the Aggregate-Distribution Tests stand apart from other existing tests by taking advantage of information about individual-level uncertainty in beliefs (and potentially other higher moments). However, the former considers the mean uncertainty of individual beliefs ($\delta_{t,1}^B$ in equation 4) whereas the latter considers the variance (and higher moments) of the aggregate distributions of beliefs and realizations. These are different but related objects. For example, the variance of the aggregate distribution of beliefs is equal to the sum of mean uncertainty of individual beliefs and the variance of individual means ($var_i(\mu_{it}^B)$). The latter is referred to as the heterogeneity of beliefs as in Gong, Stinebrickner, and Stinebrickner (2019).

As discussed in Section 4.3.1 and 4.3.2, the implementation and benefits of the Aggregate-Distribution test depend on how probabilistic expectations are collected. When a probability format is used, a non-parametric test that naturally takes advantage of all available information about the shape of individuals’ belief distributions can be formulated. When a percentile

format is used, a parametric assumption is required.

4.3.1 A Nonparametric Aggregate-Distribution Test

Beginning with the case where probabilistic expectations are elicited with the probability format, the Rational Expectations Hypothesis defined in Section 4.1 implies an equality between an individual's perceived probability that the outcome is in the set S_k , $P_{it}^{B,k}$, and the actual probability that Y_i is in S_k conditional on $X_i^{\Theta_i, t^-} = x_i^{\theta_i, t^-}$ and $Z^{t^-} = z^{t^-}$.

Corollary 1. *If individual i has Rational Expectations about Y_i (Definition 1), then*

$$P_{it}^{B,k} \equiv \text{Prob}(Y_{it}^B \in S_k) = \text{Prob}(Y_i \in S_k | (\mathbf{X}_i^{\Theta_i, t^-} = \mathbf{x}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-} = \mathbf{z}^{t^-})).$$

Corollary 1 implies the following theorem:

Theorem 3. *Under assumptions A1 to A3, if all individuals have Rational Expectations about Y_i (Definition 1), then*

$$E_i(P_{it}^{B,k}) = \text{Prob}(Y_i \in S_k | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}), \quad \text{for all } k.$$

Proof. As above, for ease of exposition, we prove the theorem for the case where Θ_i and \mathbf{Z}^{t^-} are degenerate. In this case, $\text{Prob}(Y_i \in S_k | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}) = \text{Prob}(Y_i \in S_k)$. We simplify notation by suppressing the θ_i superscript. In Appendix D we prove Theorem 3, allowing for nondegenerate Θ_i and \mathbf{Z}^{t^-} .

Let $I(\cdot)$ denote an indicator function that takes a value of one when the statement is true and takes a value of zero when the statement is false. Under the assumption that individual i has Rational Expectations about Y_i , her perceived probability $P_{it}^{B,k}$ is given by:

$$\begin{aligned} P_{it}^{B,k} &= \text{Prob}(Y_i \in S_k | \mathbf{X}_i^{t^-} = \mathbf{x}_i^{t^-}) \\ &= E[I(Y_i \in S_k) | \mathbf{X}_i^{t^-} = \mathbf{x}_i^{t^-}]. \end{aligned} \tag{12}$$

The population distribution of $P_{it}^{B,k}$ is produced by the realizations of $\mathbf{X}_i^{t^-}$. Hence,

$E_i(P_{it}^{B,k})$ is given by:

$$\begin{aligned} E_i(P_{it}^{B,k}) &= E\{E[I(Y_i \in S_k)|\mathbf{X}_i^{t-}]\} \\ &= E[I(Y_i \in S_k)] \\ &= \text{Prob}(Y_i \in S_k), \end{aligned} \tag{13}$$

where the second line directly follows from the law of iterative expectations. \square

Theorem 3 implies that, in our baseline environment where there are no aggregate shocks after time t , the test of the Rational Expectations Hypothesis is equivalent to a test for the equality between the population mean of the perceived probability, $P_{it}^{B,k}$, and the fraction of all realizations that fall in the set S_k , $\text{Prob}(Y_i \in S_k|\mathbf{Z}^{t-} = \mathbf{z}^{t-})$.

We can allow the reported perceived probability $\tilde{P}_{it}^{B,k}$ to contain classical measurement error η_i^k :

$$\tilde{P}_{it}^{B,k} = P_{it}^{B,k} + \eta_i^k, \tag{14}$$

where η_i^k has a mean of zero and is independent of $P_{it}^{B,k}$. We allow η_i^k to be arbitrarily correlated across k . Then, $E_i(P_{it}^{B,k})$ is consistently estimated by the sample mean of $\tilde{P}_{it}^{B,k}$. We let $\tilde{\mathbf{P}}_{it}^B$ denote the vector of reported perceived probabilities for person i at time t : $\tilde{\mathbf{P}}_{it}^B \equiv (\tilde{P}_{it}^{B,1}, \tilde{P}_{it}^{B,2}, \dots, \tilde{P}_{it}^{B,K-1})$.

Turning to the outcome of interest, we define Q_i^k to have a value of one if y_i is in set S_k and to have a value of zero otherwise. Let \mathbf{Q}_i denote a vector of Q_i^k : $\mathbf{Q}_i \equiv (Q_i^1, Q_i^2, \dots, Q_i^{K-1})$. Then, $\text{Prob}(Y_i \in S_k|\mathbf{Z}^{t-} = \mathbf{z}^{t-})$ is consistently estimated by the sample mean of Q_i^k .

As mentioned earlier, this test does not require panel data. In general we allow for an unbalanced panel (in which not all respondents report both beliefs and realizations), with a balanced panel and repeated cross-sections as special cases. We let \mathcal{P}_I contain $\tilde{\mathbf{P}}_{it}^B$ for all individuals who report both beliefs and realizations, and let \mathcal{Q}_I contain \mathbf{Q}_i for all individuals who report both beliefs and realizations. Similarly, let \mathcal{P}_{II} contain $\tilde{\mathbf{P}}_{it}^B$ for all individuals who only report beliefs, and we let \mathcal{Q}_{II} contain \mathbf{Q}_i for all individuals who only report realizations.

Theorem 4. (Nonparametric Aggregate-Distribution Test) Denote the mean and variance of $\tilde{\mathbf{P}}_{it}^B$ for sample $\{\mathcal{P}_I, \mathcal{P}_{II}\}$ as $\hat{\mu}_P$ and \hat{V}_P , respectively, and denote the mean and variance of \mathbf{Q}_i for sample $\{\mathcal{Q}_I, \mathcal{Q}_{II}\}$ as $\hat{\mu}_Q$ and \hat{V}_Q , respectively. Let \hat{V}_{PQ} and \hat{V}_{QP} denote the cross-covariance matrices of $\tilde{\mathbf{P}}_{it}^B$ and \mathbf{Q}_i for sample $\{(\tilde{\mathbf{P}}_{it}^B, \mathbf{Q}_i) | \tilde{\mathbf{P}}_{it}^B \in \mathcal{P}_I, \mathbf{Q}_i \in \mathcal{Q}_I\}$. Finally, we denote the number of individuals who reported both probabilities and realizations as n_I^{PQ} ,

and denote the number of individuals who are in \mathcal{P}_{II} and \mathcal{Q}_{II} , respectively, as n_{II}^P and n_{II}^Q , respectively.

Under assumptions A1 to A3, if all individuals have Rational Expectations about Y_i (Definition 1), then

$$\hat{u}_p \equiv (\hat{\mu}_P - \hat{\mu}_Q)' \hat{V}^{-1} (\hat{\mu}_P - \hat{\mu}_Q) \xrightarrow{d} \chi^2(K-1),$$

$$\text{where } \hat{V} = \frac{\hat{V}_P}{n_I^{PQ} + n_{II}^P} + \frac{\hat{V}_Q}{n_I^{PQ} + n_{II}^Q} - \frac{n_I^{PQ}}{(n_I^{PQ} + n_{II}^P)(n_I^{PQ} + n_{II}^Q)} (\hat{V}_{PQ} + \hat{V}_{QP}).$$

Theorem 4 allows us to form a test of the Rational Expectations Hypothesis based on the asymptotic right-tailed p-value associated with the test statistic \hat{u}_p . One assumption that is crucial for Theorem 4 is that data are Missing Completely at Random, which requires that the individuals for whom we observe both probabilities and realizations, the individuals for whom we only observe probabilities, and the individuals for whom we only observe realizations are all drawn from the same population distribution. We explore the sensitivity of our results to different data structures (e.g., balanced panel or repeated cross-section) in Section 5, when we apply this test to post-college income beliefs.

To summarize, this test has several appealing features. First, the test does not require parametric assumptions on the distribution of Y_{it}^B , and the test leverages all of the probability information that is directly available in the data. Second, because this test does not require a balanced panel, it is compatible with a variety of sampling schemes that are encountered in practice. This is particularly important because much research on expectations draws on repeated cross sections or even expectations and realizations recorded in different surveys of the same population Manski (2004, 2018). Third, this test is robust to classical measurement error in responses to expectations questions on surveys. In contrast, the Mean-Variance test developed in Section 4.2 will be sensitive to measurement error because, for example, the measured aggregate “actual” uncertainty $\delta_{t,2}^B = \text{var}_i(y_i - \mu_{it}^B)$ will overstate its true value if the measured value of μ_{it}^B contains classical measurement error. Ameriks et al. (2020) and Gong, Stinebrickner, and Stinebrickner (2019, forthcoming) document a non-trivial amount of measurement error in responses to survey expectations questions. However, we note again that measurement error in elicited probabilistic expectations may not be classical, for example because of rounding (Manski and Molinari, 2010).

4.3.2 A Parametric Aggregate-Distribution Test

The nonparametric test developed in Section 4.3.1 cannot be directly applied when probabilistic expectations are elicited with the percentile format. This is because (in contrast to the probability format) averaging the same percentile of Y_{it}^B across individuals does not produce the corresponding percentile of the aggregate subjective distribution. Nonetheless, we can form an aggregate-distribution test for the percentile format if we impose parametric assumptions on Y_{it}^B , which allow moments of the aggregate subjective distribution of Y_{it}^B to be obtained from the reported percentiles. Then, within our aggregate-distribution approach, we can construct a test of Rational Expectations based on moment-by-moment equalities between the aggregate subjective distribution and the distribution of realizations. We focus on the mean and variance, though in principle further moments could be incorporated into the test.

Formally, Theorem 5 shows that a parametric aggregate-distribution test can be based on the first and second moments of the aggregate subjective distribution and distribution of outcomes. The proof for Theorem 5 is generally similar to the proof for Theorem 1. We show the details in Appendix D.

Theorem 5. *Under assumptions A1 to A3, if all individuals have Rational Expectations about Y_i (Definition 1), then*

$$E_i(y_i) = E_i(\mu_{it}^B) \text{ and } E_i(\text{var}(Y_{it}^B)) + \text{var}_i(\mu_{it}^B) = \text{var}_i(y_i).$$

Since the mean of the aggregate subjective distribution is equal to the population mean of $\mu_{it}^B = E(Y_{it}^B)$, an aggregate-distribution test of Rational Expectations based on the equality between the means of the aggregate subjective distribution and the aggregate distribution of realizations ($E_i(y_i) = E_i(\mu_{it}^B)$) is essentially the mean component of the Mean-Variance Test ($\sqrt{\delta_{t,3}^B} = 0$) presented in Section 4.2. The only difference is that, while the mean component of the Mean-Variance Test in Section 4.2 requires balanced panel data, here the mean component of the Aggregate-Distribution Test can be applied to unbalanced panels and repeated cross-sections from a common population. We let $\hat{\mu}_{BM}$ and $\hat{\sigma}_{BM}$ respectively denote the sample mean and standard deviation of μ_{it}^B for all individuals for whom we observe Y_{it}^B , and let n^B denote the size of this sample. Similarly, we let $\hat{\mu}_{YM}$ and $\hat{\sigma}_{YM}$ respectively denote the sample mean and standard deviation of y_i for all individuals for whom we observe y_i , and let n^Y denote the size of this sample. Finally, we let $\hat{\sigma}_M^2$ denote the sample covariance

of μ_{it}^B and y_i for all individuals for whom we observe both Y_{it}^B and y_i , and let n^{BY} denote the size of this sample.

On the other hand, as noted above, the variance of the aggregate belief distribution is the sum of mean individual uncertainty and the variance of individual mean beliefs (heterogeneity in the terminology of Gong, Stinebrickner, and Stinebrickner, 2019). This means that the variance component of the Aggregate-Distribution Test differs from the test presented in Section 4.2. Nevertheless, $E_i(\mu_{it}^B)$, $E_i(\text{var}(Y_{it}^B))$, and $\text{var}_i(\mu_{it}^B)$ can all be constructed from cross-sectional empirical distributions under a parametric assumption on Y_{it}^B . We let $\hat{\mu}_{BV}$ and $\hat{\sigma}_{BV}$ denote the sample mean and standard deviation of $\text{var}(Y_{it}^B) + (\mu_{it}^B - \hat{\mu}_{BM})^2$ for all individuals for whom we observe Y_{it}^B , respectively. Similarly, we let $\hat{\mu}_{YV}$ and $\hat{\sigma}_{YV}$ denote the mean and standard deviation of $(y_i - \hat{\mu}_{YM})^2$ for all individuals for whom we observe y_i . Finally, we let $\hat{\sigma}_V^2$ denote the covariance of $\text{var}(Y_{it}^B) + (\mu_{it}^B - \hat{\mu}_{BM})^2$ and $(y_i - \hat{\mu}_{YM})^2$ for all individuals for whom we observe both Y_{it}^B and y_i .

Under the following assumptions on the distribution of random variables, we are able to define and characterize the asymptotic distribution of two test-statistics using the central limit theorem.

A5 The following random variables have finite second moments: μ_{it}^B , y_i , $\text{var}(Y_{it}^B) + (\mu_{it}^B - \mu_t^B)^2$, $(y_i - \bar{y})^2$, $\mu_{it}^B - y_i$, and $\text{var}(Y_{it}^B) + (\mu_{it}^B - \mu_t^B)^2 - (y_i - \bar{y})^2$.

Theorem 6. (*Parametric Aggregate-Distribution Test*) Under assumptions A1 to A3 and A5, if all individuals have Rational Expectations about Y_i (Definition 1), then

$$\hat{u}_M \equiv \frac{\hat{\mu}_{BM} - \hat{\mu}_{YM}}{\sqrt{\frac{\hat{\sigma}_{BM}^2}{n^B} + \frac{\hat{\sigma}_{YM}^2}{n^Y} - \frac{2n^{BY}\hat{\sigma}_M^2}{n^B n^Y}}} \xrightarrow{d} N(0, 1),$$

$$\hat{u}_V \equiv \frac{\hat{\mu}_{BV} - \hat{\mu}_{YV}}{\sqrt{\frac{\hat{\sigma}_{BV}^2}{n^B} + \frac{\hat{\sigma}_{YV}^2}{n^Y} - \frac{2n^{BY}\hat{\sigma}_V^2}{n^B n^Y}}} \xrightarrow{d} N(0, 1).$$

Theorem 6 suggests that two different Aggregate-Distribution Tests of Rational Expectations can be formed based on the asymptotic two-tailed p-values of the test statistic \hat{u}_M and \hat{u}_V , respectively.

This version of the Aggregate-Distribution Test follows the same general approach as the nonparametric version and retains the advantages of accommodating a wide range of data structures (panels, repeated cross-sections) and incorporating higher-moment information

about belief distributions. Moreover, as we show in Section 6, it can be adjusted to accommodate the possibility of aggregate shocks in some contexts where the nonparametric version cannot. However, unlike the nonparametric version, it does require a parametric assumption about belief distributions. One virtue of the nonparametric version of the Aggregate-Distribution test, applied to beliefs elicited in the probability format, is robustness to classical measurement error in the reporting of probabilities. It turns out that if the parametric version of the Aggregate-Distribution test is applied to beliefs elicited in the probability format and if individual belief distributions are such that the aggregate belief distribution can be calculated from average perceived probabilities, then the parametric version of the Aggregate-Distribution test is also robust to classical measurement error in the reporting of probabilities in the same way. The reason is that, in this case, the test is identical to one directly calculated from movements of the aggregate distribution, which average out errors across respondents. The step-wise uniform assumption on individual belief distributions that we maintain throughout for empirical application below satisfies this condition, but it is not the only possibility.⁸

4.4 Existing Mean-based Tests of Rational Expectations

Most existing tests of Rational Expectations utilize only information about the mean of Y_{it}^B , μ_{it}^B . It is important to recognize that existing mean-based tests potentially detect different types of violations of Rational Expectations than our new tests do. We now compare existing mean-based tests to our tests, continuing, for the time being, to focus on the scenario where there is no possibility of aggregate shocks between the time when beliefs are elicited and the time when outcomes are realized (i.e., Z^{t+} is degenerate).

Repeated cross-section mean-based tests are based on whether moments of the distribution of μ_{it}^B and moments of the distribution of y_i satisfy certain equalities and inequalities implied by Rational Expectations. For example, in the absence of aggregate shocks, the sample mean of μ_{it}^B is asymptotically identical to the sample mean of y_i . Similarly, due to the presence of factors realized after t , X_i^{t+} , the sample variance of μ_{it}^B is asymptotically weakly smaller than the sample variance of y_i . A repeated cross-section mean-based test that takes into account all moments of μ_{it}^B and y_i is developed in D'Haultfoeuille, Gaillac, and Maurel (2018). It is based on the insight that the population distribution of y_i is a mean-preserving spread of

⁸A sufficient condition is that, within each bin, the *conditional* density function is homogeneous across agents.

the population distribution of μ_{it}^B . We refer to this test as the DGM test.

Panel mean-based tests are based on the relationships between μ_{it}^B and y_i (and possibly also \mathbf{X}_i^{t-}) that should be satisfied on an individual level under Rational Expectations (Definition 1). Rational Expectations implies y_i is the sum of μ_{it}^B and a residual term, which is uncorrelated with μ_{it}^B . A common approach to testing this assumption is to regress y_i on μ_{it}^B , and to test whether the slope coefficient is equal to 1 and whether the intercept is equal to 0. We refer to this test as the *Regression Test*. A concern with the Regression Test, when applied to micro data, is that it is potentially very sensitive to reporting error in beliefs or to approximation error in the imputation of mean beliefs μ_{it}^B . Measurement error in μ_{it}^B can lead to attenuation of the estimated slope parameter, and spurious rejection of a true null hypothesis.

A stronger version of Rational Expectations assumes that agents use all available information in their forecasts. This is often assessed by testing whether the correlation between the innovation (the difference between the realization y_i and the expectation μ_{it}^B) is orthogonal to factors that should be in respondents' information sets at the time when beliefs are elicited. We refer to the test built on this condition as the *Correlation Test*.

It is useful to explore the relationship between the mean-based tests of Rational Expectations (DGM, Regression) and our new tests (Mean-Variance, Aggregate-Distribution). With respect to the DGM Test, we note that, in the limiting case where we have full information about Y_{it}^B (i.e., infinitely many percentiles or probabilities), the DGM test is weaker than our new Aggregate-Distribution Test in the absence of aggregate shocks. This is the case because, by construction, the aggregate subjective distribution is a mean-preserving spread of the (cross-sectional) distribution of μ_{it}^B . Hence, an equality between the aggregate subjective distribution and the distribution of y_i implies that the latter is also a mean-preserving spread of the distribution of μ_{it}^B . However, since the number of bins or percentiles collected in a survey question must be modest in practice, we stress that the DGM Test and our Aggregate-Distribution Tests can both be useful.

With respect to the Regression Test, we note that there is a tight link among (1) the Regression Test, (2) the Mean-Variance Test, and (3) the Parametric Aggregate-Distribution Test (based on the mean and variance). Specifically, the following theorem holds:⁹

Theorem 7. *Any two of the following three statements imply the third statement:*

(1) $\delta_{t,1}^B = \delta_{t,2}^B$, and $\sqrt{\delta_{t,3}^B} = 0$; (*Mean-Variance Test*)

⁹See Appendix D for the proof.

- (2) $E_i(\text{var}(Y_{it}^B)) + \text{var}_i(\mu_{it}^B) = \text{var}_i(y_i)$, and $E_i(y_i) = E_i(\mu_{it}^B)$; (*Aggregate-Distribution Test*)
(3) $y_i = \mu_{it}^B + \epsilon_i$, where ϵ_i has a mean of zero and is orthogonal to μ_{it}^B . (*Regression Test*)

This suggests that a joint test combining the Mean-Variance and Parametric Aggregate-Distribution Test may be useful. As the mean component of these two tests is identical, the joint test is a test of three moment conditions. This combination of these two tests provides a broad assessment of the Rational Expectations Hypothesis, which encompasses the well-known Regression Test, but also provides information beyond that test. In other words, asymptotically, the joint test will reject Rational Expectations whenever the Regression Test rejects, but not vice versa.

5 Testing Rational Expectations: Empirical Results

We now turn to testing Rational Expectations about future incomes with data from the BPS. It is worth noting that, in this section, we are in effect testing a joint null that all respondents have Rational Expectations about future income and that there is no possibility of aggregate shocks between the time beliefs are elicited and the time outcomes are realized. However, our interpretation of the results abstracts from the possibility of aggregate shocks. We then explicitly consider issues related to this possibility in the next Section (Section 6).

One goal of this section is to examine whether Rational Expectations provides a better description of beliefs in certain periods than in other periods. In Section 5.1 we focus on a comparison of in-school and post-college beliefs by using our new tests to examine the rationality of beliefs about future income at the fixed age of 28 (or 29). A second goal of this section is to examine the practical importance of our new tests relative to existing mean-based tests. While theoretical results in Section 4 show that our new tests, which take into account higher moments, are able to detect more types of deviations from Rational Expectations, exactly how important this is in practice is an open empirical question. To provide a comparison of our tests and existing tests, we take advantage of the 1-year ahead and 5-year ahead beliefs that were collected in the post-college period. The fact that these beliefs were collected annually for an extended period of time provides a substantial number of observations for our comparison. The fact that these beliefs were collected using a probability format allows us to implement our Nonparametric Aggregate-Distribution test. Section 5.2 also complements the fixed age analysis in Section 5.1 by examining how the rationality of beliefs about income changes over time when the time horizon is fixed.

5.1 A Comparison of In-school and Post-College Beliefs

Mean-Variance Test

To compare the rationality of in-school beliefs to the rationality of post-college beliefs, we begin by utilizing the Mean-Variance test to study beliefs about income at the fixed age of 28 (or 29). Given the link between this test and measures of ex post inaccuracy, an indication of what we will find can be obtained by referring back to Table 2 in Section 3. The third column shows that $\delta_{t,3}^B$ is decreasing over time, indicating declining mean bias. Moreover, mean subjective uncertainty ($\delta_{t,1}^B$) and the variance of income innovations ($\delta_{t,2}^B$) are more similar in the post college period. Hence, Table 2 suggests that Rational Expectations might provide a more apt characterization of beliefs about income during the post-college period than during the college period.

Table 3 reports the p-values associated with the two components of our (two-tailed) Mean-Variance test. Panel A shows results for the in-school period. The first row of Panel A shows that the null hypothesis that $\delta_{t,1}^B = \delta_{t,2}^B$, i.e., that the average subjective uncertainty is equal to the aggregate actual uncertainty, is rejected at a 1% level for all in-school years (10-year ahead to 6-year ahead). The second row of Panel A shows that the null hypothesis that $\sqrt{\delta_{t,3}^B} = 0$, i.e., that there is no mean bias, is also rejected at a 1% level for all in-school periods. Hence, these findings indicate that, when the students are in college, their beliefs about future income deviate non-trivially from Rational Expectations (assuming no aggregate shocks). They suffer from mean bias and also insufficient subjective uncertainty. Panel B shows results for the post-college period. In contrast to the findings for the in-school period in Panel A, Panel B shows that Rational Expectations is frequently not rejected during the post-college period. As shown in the first row of Panel B, while we reject the null hypothesis that $\delta_{t,1}^B = \delta_{t,2}^B$ at a 1% level for 5-year ahead income beliefs, we do not reject this null hypothesis even at a 5% level for 1-year ahead income beliefs. The second row of Panel B shows that we do not reject the null hypothesis that $\sqrt{\delta_{t,3}^B} = 0$ at a 5% level for either 1-year ahead or 5-year ahead post-college beliefs about income at age 28 (29).

The Parametric Aggregate-Distribution Test

Turning to using the parametric Aggregate-Distribution Test to examine beliefs at the age of 28 (29), we are not limited to the balanced panel, so Table 4 presents results both for the balanced panel and for the pooled data (all observations on beliefs and realizations for a given year). The results are very similar to what was found for the Mean-Variance

test. Panel A shows that, regardless of which set of observations is used, for all in-school years we reject at a 1% level the null hypothesis that the mean of the subjective aggregate distribution is equal to the mean of the distribution of outcomes. Similarly, Panel B shows that, except for the final year of school (6-year ahead beliefs), we also reject the variance component of the null hypothesis at a 1% level for both sets of observations; in this final year, we reject this null hypothesis at a 1% level for the balanced panel, but not for the pooled data. The last two columns of Table 4 show that, as was the case with the Mean-Variance test, we reject Rational Expectations less often in the post-college period. Panel A shows that, for both 1-year ahead beliefs and 5-year ahead beliefs, the mean component of the Aggregate-Distribution Test does not reject at a 5% level using either the balanced panel or the pooled data. Panel B shows that, for both the 1-year ahead beliefs and 5-year ahead beliefs, the variance component rejects at a 5% level for the pooled data but not for the balanced panel.

In addition to results shown in Table 3 and 4, we also implemented a joint test of the combined elements of the Mean-Variance and Parametric Aggregate-Distribution test, which was proposed at the end of Section 4. The joint test rejects at a 1% level for all in-school beliefs and for 5-year ahead post-college beliefs. However, it does not reject at a 10% level for 1-year ahead post-college beliefs (with a p-value of 0.2171).

Under the assumption that there is no possibility of aggregate shocks between belief elicitation and income realization, the evidence against the rationality of in-school beliefs is very strong. In contrast, the evidence for the post-college period is weaker; we reject the joint null for some tests at some horizons, but not others. In Appendix E, we show that this general conclusion remains when we employ existing mean-based tests.

5.2 Post-College by Income Realization Year

The 1-year and 5-year ahead beliefs collected annually in the post-college period of the BPS provide ten belief-realization combinations; 1-year ahead beliefs can be compared to realizations from each of the years between 2007-2013, and 5-year ahead beliefs can be compared to realizations from each of the years between 2011-2013. This relatively large number of belief-realization combinations is beneficial for comparing the performance of our tests to the performance of existing mean-based tests. In particular we are interested in examining the practical relevance of theoretical results from Section 4 that describe the relationship between: 1) our Nonparametric Aggregate-Distribution test and the DGM test

Table 3: Mean-Variance Test about Income at Age 28-29 (p-values)

Year of Belief-elicitation	Panel A					Panel B	
	10-year ahead	9-year ahead	8-year ahead	7-year ahead	6-year ahead	5-year ahead	1-year ahead
\hat{u}_{12} (uncertainty)	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	0.0687
\hat{u}_3 (mean)	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	0.0762	0.7077
Sample Size	271	259	206	202	207	159	289

Notes: Panel A reports the p-values associated with (two-tailed) test statistics \hat{u}_{12} and \hat{u}_3 (defined in Theorem 2) for students' beliefs in college about own annual income at age 28 (or 29). Panel B reports the p-values associated with test statistics \hat{u}_{12} and \hat{u}_3 for students' beliefs after college about annual family income at age 28 (or 29).

Table 4: Parametric Aggregate-Distribution Test about Income at Age 28-29 (p-values)

Year of Belief-elicitation	10-year ahead	9-year ahead	8-year ahead	7-year ahead	6-year ahead	5-year ahead	1-year ahead
Panel A: \hat{u}_M (p-value)							
All Observations	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	0.9797	0.2790
Common Observations	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	0.0762	0.7077
Panel B: \hat{u}_V (p-value)							
All Observations: all	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	0.0035	0.0029	0.0013
Common Observations: common	$< 10^{-4}$	$< 10^{-4}$	0.0001	0.0039	0.2079	0.3030	0.0797
Panel C: Sample Size							
n^{BY}	271	259	206	202	207	159	289
n^B	667	581	414	371	361	214	339
n^Y	330	330	330	330	330	375	360

Notes: The first five columns of Panel A and Panel B report the p-values associated with (two-tailed) test statistics \hat{u}_M and \hat{u}_V (defined in Theorem 6) for students' beliefs in college about own annual income at age 28 (or 29). The last two columns of Panel A and Panel B report the p-values associated with test statistics \hat{u}_M and \hat{u}_V for students' beliefs after college about annual family income at age 28 (or 29).

and 2) our joint test (of the Mean-Variance and Parametric Aggregate-Distribution tests) and the Regression test. The belief-realization combinations can also be used to examine the rationality of beliefs for different realization years while holding the time horizon fixed.

A Comparison of the Nonparametric Aggregate-Distribution and DGM Tests

In Section 4 we explained why, in the limiting case where a large number of bins/probabilities are available, the DGM test is weaker than our new Nonparametric Aggregate-Distribution Test in the absence of aggregate shocks. To examine the practical importance of this result, we apply these two tests to the BPS data and compare the results.

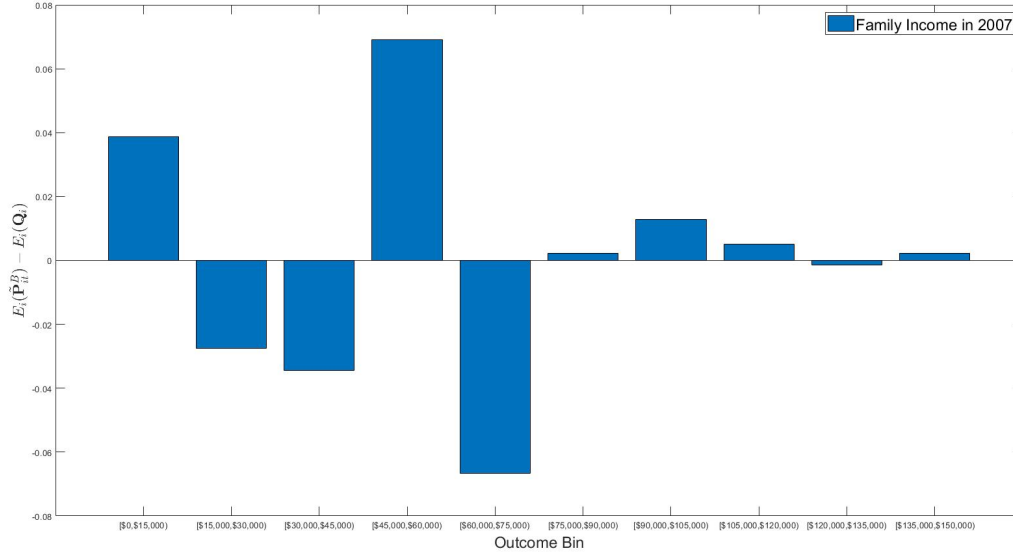
We begin by implementing the DGM test. We consider two sets of observations, the pooled data (which exploits the fact that this test can be implemented with repeated cross-sections or unbalanced panels) and the balanced panel. We find that, regardless of which set of observations we use, for each of the ten belief-realization combinations, we do not reject the null hypothesis of Rational Expectations at any conventional significance levels (the smallest p-value is 0.18 for the pooled data and 0.21 for the balanced panel).¹⁰

We next consider the Nonparametric Aggregate-Distribution Test, which fully exploits the probability format of the belief questions in the post-college surveys. As described in Section 4.3.1, for each income interval defined by the survey question, this test compares the average (across respondents) subjective probability reported for the interval to the proportion of realized incomes that fall in the interval. To illustrate, the difference in these quantities is shown in Figure 1 for 1-year ahead beliefs about family income in 2007 and 2008. Then, because the differences are in general much smaller in 2008 (Panel (b)), we might expect the test to be less likely to reject in this year.

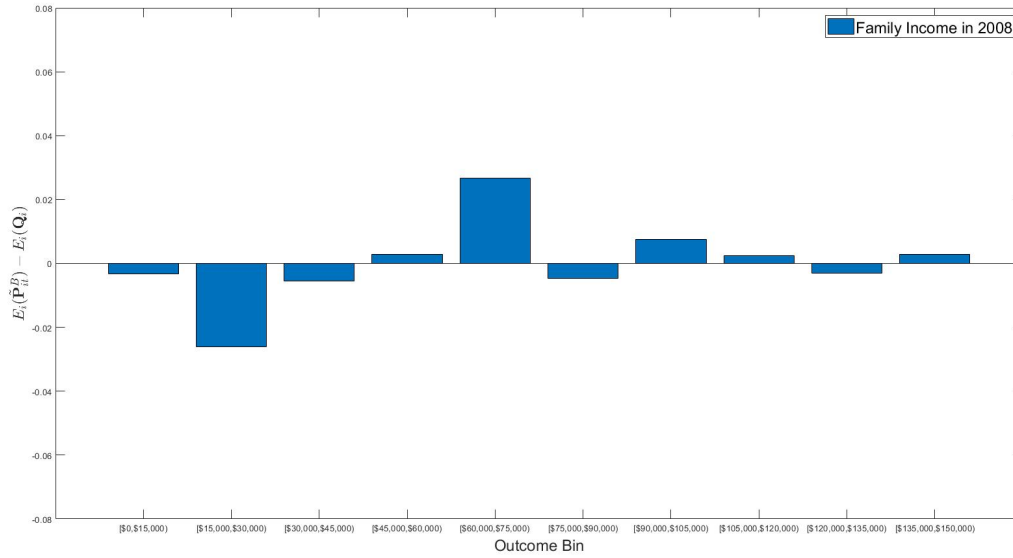
Table 5 reports the results of the tests for all belief-realization combinations, where again we examine results for both the pooled data and the balanced panel. Panel A summarizes the results for 1-year ahead beliefs, with asymptotic p-values reported in Panel A2. We find that, regardless of which set of observations we use, we reject at a 5% level the null hypothesis of Rational Expectations about 1-year ahead family income in 2007, 2009, and 2012, so that the conclusions about 2007 and 2008 are consistent with our discussion of Figure 1. Panel B shows that we also reject at a 5% level the null hypothesis of Rational Expectations about 5-year ahead annual family income in 2012 and 2013.

Thus, the hypothesis of Rational Expectations is rejected for five of the ten belief-

¹⁰We implement the DGM test using the R-package provided by the authors of D'Haultfoeuille, Gaillac, and Maurel (2018).



(a) Family Income Realized in 2007



(b) Family Income Realized in 2008

Figure 1: Differences between Average Perceived and Actual Probability for 10 Income Bins

Table 5: Nonparametric Aggregate-Distribution Test about Income in 2007-2013 (p-values)

Year of Realization	2007	2008	2009	2010	2011	2012	2013
Panel A: 1-year ahead Beliefs							
Panel A1: Sample Size							
\mathcal{P}_I and \mathcal{Q}_I	202	262	272	296	322	327	346
\mathcal{P}_{II}	42	53	60	42	39	46	34
\mathcal{Q}_{II}	119	84	70	68	51	51	50
Panel A2: p-value							
$\{\mathcal{P}_I, \mathcal{P}_{II}, \mathcal{Q}_I, \mathcal{Q}_{II}\}$: pooled data	0.0242	0.6700	$< 10^{-4}$	0.0840	0.6139	0.0114	0.5580
$\{\mathcal{P}_I, \mathcal{Q}_I\}$: balanced panel	0.0002	0.1498	$< 10^{-4}$	0.3691	0.3235	0.0015	0.0914
Panel B: 5-year ahead Beliefs							
Panel B1: Sample Size							
\mathcal{P}_I and \mathcal{Q}_I	N.A.	N.A.	N.A.	N.A.	189	240	266
\mathcal{P}_{II}	N.A.	N.A.	N.A.	N.A.	58	74	70
\mathcal{Q}_{II}	N.A.	N.A.	N.A.	N.A.	184	138	130
Panel B2: p-value							
$\{\mathcal{P}_I, \mathcal{P}_{II}, \mathcal{Q}_I, \mathcal{Q}_{II}\}$: pooled data	N.A.	N.A.	N.A.	N.A.	0.0612	0.0098	0.0006
$\{\mathcal{P}_I, \mathcal{Q}_I\}$: balanced panel	N.A.	N.A.	N.A.	N.A.	0.0808	0.0167	0.0010

Notes: Panel A reports the p-values associated with (two-tailed) test statistics \hat{u}_p (defined in Theorem 4) for students' 1-year ahead beliefs after college about annual family income realized in 2007-2013. Panel B reports the p-values associated with (two-tailed) test statistics \hat{u}_p for students' 5-year ahead beliefs after college about annual family income realized in 2011-2013.

realization combinations when the Nonparametric Aggregate-Distribution test is employed, but is never rejected when the DGM test is employed. This suggests that, in practice, the consideration of higher moments may be quite important for characterizing whether beliefs are rational.

Comparing Our Joint Test with the Regression Test

As stated in Theorem 7, the null hypotheses of the Mean-Variance and Parametric Aggregate-Distribution Tests jointly imply the null of the Regression Test (but not vice versa), suggesting that the Joint Test might be more likely to reject than the Regression Test. To examine the practical importance of this result, we apply these two tests (and the individual components of these tests) to the BPS data and compare the results. Because the Regression and Mean-Variance tests require a balanced panel, we focus here on that data structure.

We begin by implementing the Regression test. We regress realized annual family income on the mean of the respondent's subjective income distribution in an earlier year (either 1 year prior or 5 years prior), and jointly test whether the resulting intercept is equal to zero and whether the slope is equal to one. The results are summarized in Panel A1 and B1 of Table 6. Among the seven 1-year ahead belief-realization combinations, we reject at a 5% level in 2007, 2008, 2009, 2012, and 2013. For the 5-year ahead beliefs, the regression test rejects in 2011 and 2012.

We next consider the Joint Test (of the Mean-Variance and Parametric Aggregate-Distribution Tests). The results are shown in the second-to-last row of Panel A and the second-to-last row of Panel B of Table 7. Our theoretical result suggests that we are likely to reject the null for all of the belief-realization combinations for which the Regression test rejected (1-year ahead beliefs for 2007, 2008, 2009, 2012, and 2013 and 5-year ahead beliefs for 2011 and 2012). We find that this is the case. The theoretical results also suggests that we might find additional rejections. We find this to be true for one belief-realization combination: 5-year ahead beliefs for 2013.

To provide a more in-depth view of our joint test, Figure 2 (for 1-year ahead beliefs) and Figure 3 (for 5-year ahead beliefs) plot the components of the joint test by year. Recall that the Mean-Variance test has two elements: the cross-sectional average of mean beliefs should equal average outcomes ($\sqrt{\delta_{t,3}^B} = 0$, no mean bias) and the cross-sectional average of subjective uncertainty ($\delta_{t,1}^B$) should equal the variance of innovations ($\delta_{t,2}^B$). These four objects are plotted in Panels (a) and (b) of Figures 2 and 3. The mean component of the

Table 6: Regression and Correlation Tests of RE about Income in Year 2007-2013

Year of Realization	2007	2008	2009	2010	2011	2012	2013
Panel A: 1-year ahead Beliefs							
Panel A1: Regression Test							
Intercept	\$8,610	\$6,180	\$4,070	\$2,080	\$2,750	\$3,360	-\$4,960
Slope	0.7724	0.8078	0.8818	0.9647	0.9257	0.9957	1.1587
p-value	0.0002	0.0004	0.0213	0.5514	0.1021	0.0002	0.0001
Sample Size	200	264	269	294	320	322	344
Panel A2: Correlation Test							
Correlation	0.0372	0.1358	0.0384	0.1228	0.0425	0.0760	0.2701
p-value	0.6073	0.0290	0.5296	0.0371	0.4552	0.1735	$< 10^{-4}$
Sample Size	200	259	268	290	319	319	336
Panel B: 5-year ahead Beliefs							
Panel B1: Regression Test							
Intercept	N.A.	N.A.	N.A.	N.A.	\$14,170	\$20,730	\$6,130
Slope	N.A.	N.A.	N.A.	N.A.	0.6106	0.5469	0.8709
p-value	N.A.	N.A.	N.A.	N.A.	$< 10^{-4}$	$< 10^{-4}$	0.2981
Sample Size	N.A.	N.A.	N.A.	N.A.	184	237	265
Panel B2: Correlation Test							
Correlation	N.A.	N.A.	N.A.	N.A.	-0.0033	0.0266	0.0762
p-value	N.A.	N.A.	N.A.	N.A.	0.9643	0.6874	0.2139
Sample Size	N.A.	N.A.	N.A.	N.A.	183	232	264

Notes: Panel A1 and A2 report the results for the Regression Test (estimated parameters and p-value) and the Correlation Test (estimated correlation and p-value) for students' 1-year ahead beliefs after college about annual family income realized in 2007-2013, respectively. Panel B1 and B2 report the results for the Regression Test and the Correlation Test for students' 5-year ahead beliefs after college about annual family income realized in 2011-2013, respectively.

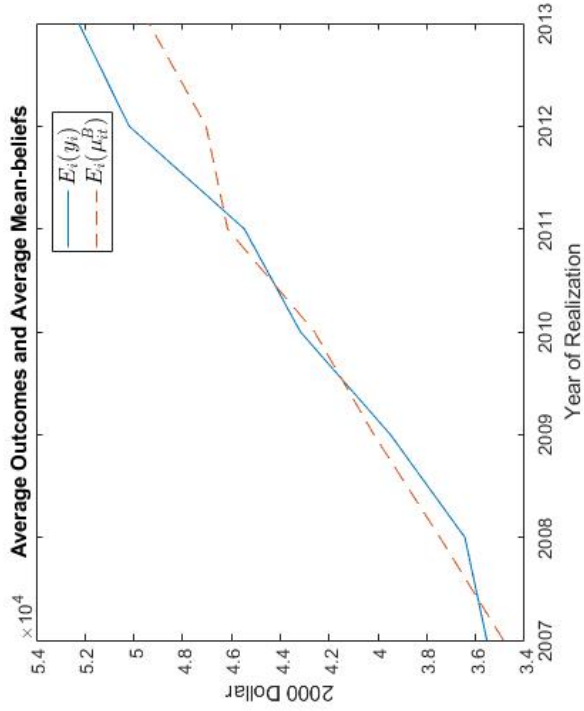
Parametric Aggregate-Distribution Test is identical to the mean component of the Mean-Variance Test when we implement our tests on balanced panels. The variance component of the Parametric Aggregate-Distribution Test compares the variance of realized income and the variance of the aggregate distribution of beliefs. These two objects are plotted in Panel (c) of Figures 2 and 3.

Looking at the 1-year ahead beliefs in Figure 2, Panel (a) shows that average mean beliefs (solid line) track average outcomes (dashed line) quite closely up until 2011, but respondents became too pessimistic about their income for 2012 and 2013. Panel (b) shows that, for every year, the variance of innovations (dashed line) exceeds the cross-sectional average of subjective uncertainty (solid line), suggesting that on average respondents underestimate the uncertainty they face. Panel (c) shows that the variance of the aggregate belief distribution (dashed line) closely tracks the variance of realized incomes (solid line), except for the very beginning and end of the sample period. These patterns are reflected in the p-values associated with tests of the individual components reported in the first three rows of Panel A in Table 7.

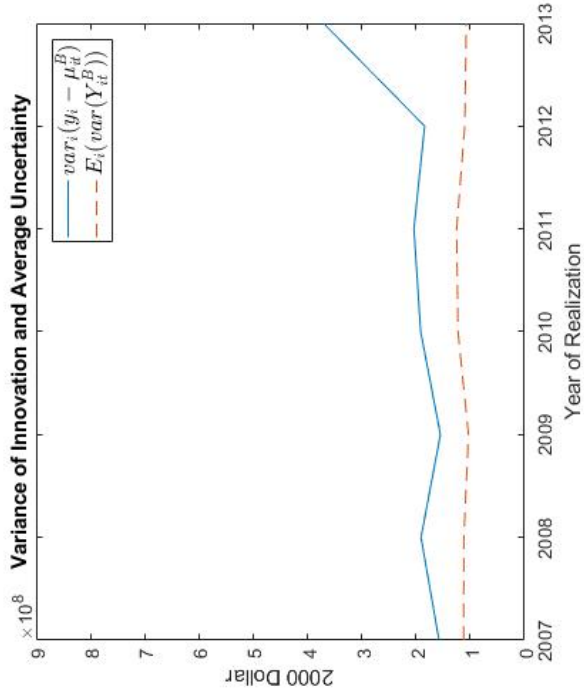
In terms of examining 5-year ahead beliefs shown in Figure 3, we focus on the year, 2013, when our Joint test produces a different result than the Regression test. The fact that the Regression test does not produce a rejection in this year is not surprising given that Panel (a) shows that the means are quite similar in this year. Indeed, the p-value in the second row of Panel B of Table 7 is 0.712 for 2013. The joint test differs primarily because it considers the variance components shown in Panel (b) and (c). Panel (b) shows that respondents vastly underestimate the amount of uncertainty associated with income for 2013. Panel (c) shows that the variance of the aggregate subjective belief distribution is much smaller than the variance of realized income for 2013. The statistical importance of these differences can be seen in Row 1 of Panel B of Table 7 (p-value .0002) and Row 3 of Panel B of Table 7 (.019). As seen in the second-to-last row of Panel B of Table 7, the variance components lead the joint test to reject with a p-value smaller than .0001. This provides a concrete example of the value of collecting and utilizing information beyond mean beliefs.

Correlation Test

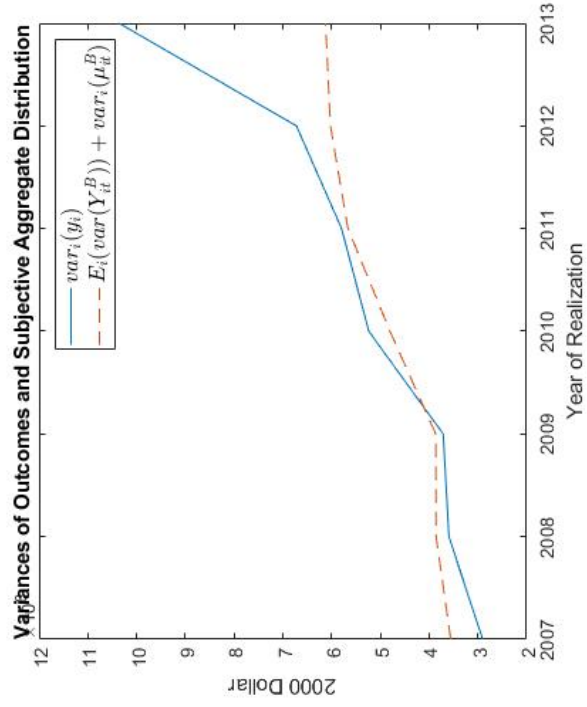
Finally, for completeness we consider the Correlation Test, even though, unlike other tests, it considers a strong version of Rational Expectations. For each year in the post-college period, we first compute the innovation, which is the difference between realized annual family income and the mean of beliefs about this income elicited 1 year or 5 years



(a) Average Outcomes and Average Mean-beliefs

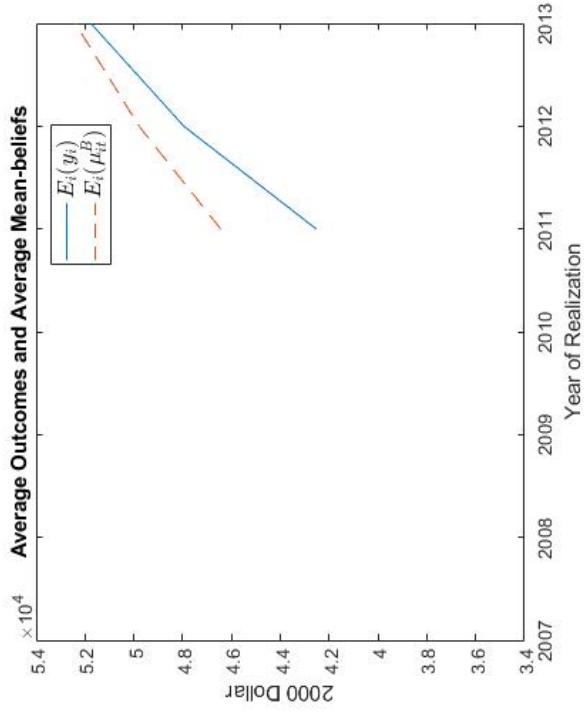


(b) Variance of Innovation and Average Uncertainty

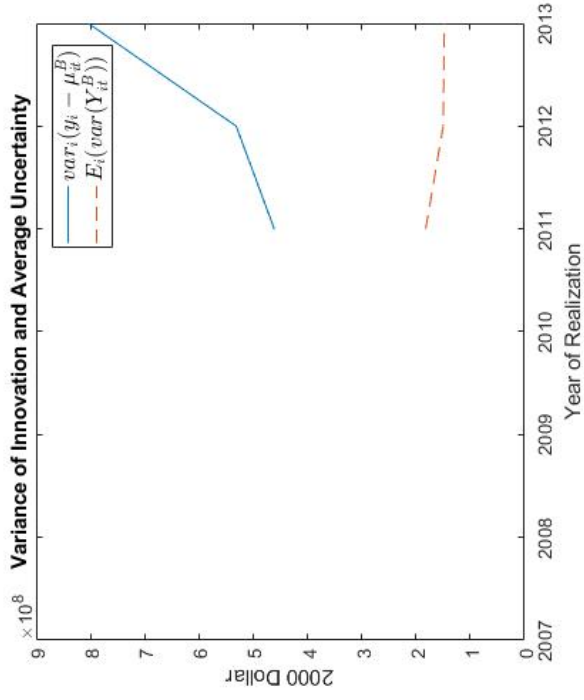


(c) Variance of Outcomes and Subjective Aggregate Distribution

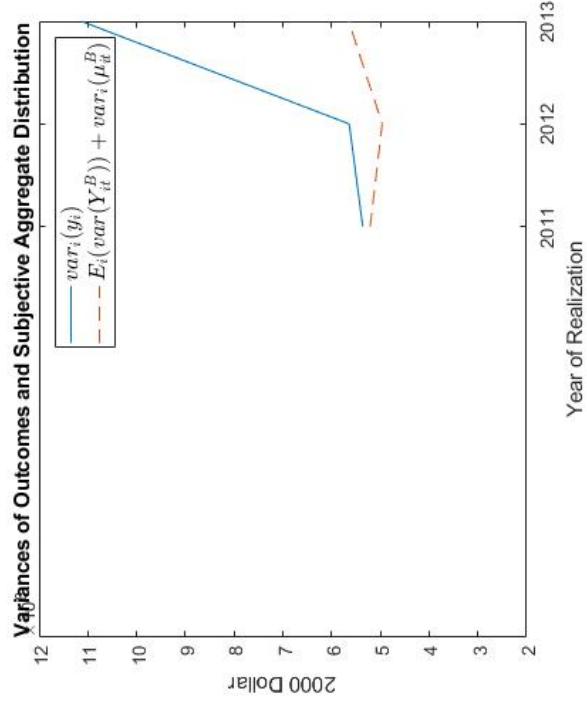
Figure 2: Descriptive Statistics (1-Year ahead Beliefs)



(a) Average Outcomes and Average Mean-beliefs



(b) Variance of Innovation and Average Uncertainty



(c) Variance of Outcomes and Subjective Aggregate Distribution

Figure 3: Descriptive Statistics (5-Year ahead Beliefs)

earlier. We then compute the cross-sectional correlation between this innovation and the realized annual family income in the year when beliefs were elicited, and test whether this correlation is equal to zero. Panel A2 of Table 6 shows the results for 1-year ahead beliefs. Across the seven post-college years, the null hypothesis that the correlation is zero is rejected at a 1% level in 2013 and is rejected at a 5% level in two other years (2008 and 2010). Our results suggest that respondents were not always using all the available information to predict their future family income.

As shown in Panel B2 of Table 6, the results for 5-year ahead beliefs are qualitatively different. For all three years that we can perform the test for 5-year ahead beliefs - 2011, 2012, and 2013 - the absolute value of point estimates of the correlation are smaller than 0.1 and we cannot reject the null hypothesis of zero correlation at a 10% level. One possible explanation of this finding is that, while current family income is highly informative about family income in the next year, it is less informative about family income in five years. Hence, a Correlation Test based on current family income may have less power when the objects of interest are 5-year ahead beliefs.

Summary

Generally, the results in Section 5 show that 1) Rational Expectations provides a better description of income beliefs for the post-college periods than the in-school periods, 2) our new tests, which take into account higher moments, reject more often than existing mean-based tests, which illustrates the practical importance of theoretical results in Section 4, and 3) across the tests that we consider, we consistently do not reject Rational Expectations in some of the post-college years, but we do not find evidence that respondents' post-college income beliefs are becoming more rational in a monotonic fashion. One potential reason for finding 3) is that, while we are in effect testing a joint null of Rational Expectations and no aggregate shocks, the interpretation given in this section abstracts away from the possibility of aggregate shocks. In practice, these shocks may be more relevant in some years than others. For example, it could be that aggregate shocks are always possible between elicitation and realization (violating the second part of the joint null) but when the realized aggregate shock is "small", then a given test does not reject in a finite sample.¹¹ However,

¹¹Asymptotically, these tests will reject a false joint null for any realization of the aggregate shock \mathbf{Z}^{t+} . An intuition is as follows. Suppose the realization of income is additive in \mathbf{Z}^{t+} and the realization of \mathbf{Z}^{t+} is zero. In this case, while the realization of \mathbf{Z}^{t+} does not cause mean bias, the ex ante possibility of other values \mathbf{Z}^{t+} will have the consequence that the distribution of beliefs is more disperse than distribution of realizations and, asymptotically, the test will detect this.

it is also possible that the conditional distributions of factors unobserved at time t (\mathbf{X}^{t+} and \mathbf{Z}^{t+}) vary over time, so that that joint null more accurately describes the environment in some years than others.¹² To explore these issues further, in Section 6 we consider Rational Expectations tests which are robust to aggregate shocks.

6 Aggregate Shocks

In previous sections, our theoretical and empirical results pertain to tests of Rational Expectations when there is no possibility of observing aggregate shocks between the time when beliefs are elicited (t) and the time when outcomes are realized. In this section, we consider how tests can be extended to accommodate the possibility of aggregate shocks.

As with most existing work on testing Rational Expectations, the tests considered in this paper utilize data on realizations from a single period. However, in this case, if no restrictions are imposed on the structure of aggregate shocks, it is always possible to find a distribution for the aggregate shock and a realization from that distribution that would imply observed beliefs are rational. In other words, without some type of restriction, it is not possible to develop a test of Rational Expectations that has power. Acknowledging this reality, we focus on a case where aggregate shocks are additively separable from and independent of other factors. Formally, we express Y_i as the sum of two components, $Y_i^I \equiv Y_I(\mathbf{X}_i)$ and $Y^A \equiv Y_A(\mathbf{Z})$, which are functions of individual-specific factors \mathbf{X}_i and aggregate factor \mathbf{Z} , respectively. Finally, consistent with the notion that \mathbf{Z} is common to all individuals, we assume that respondents' information sets with respect to aggregate factors \mathbf{Z} at the time of belief elicitation, \mathbf{Z}^{t-} , are homogeneous. This naturally implies that \mathbf{Z}^{t+} , the set of unobserved (at t) aggregate factors, is identical across people as well. We note that other researchers have imposed similar homogeneity assumptions when testing Rational Expectations in the presence of aggregate shocks (e.g., D'Haultfoeuille, Gaillac, and Maurel, 2018).

Formally, we impose the following assumptions throughout Section 6.

A6 $Y_i = Y_i^I + Y^A$.

A7 Θ_i , \mathbf{X}_i and \mathbf{Z} are independent.

A8 \mathbf{Z}^{t-} and \mathbf{Z}^{t+} are homogeneous across people.

¹²This is consistent with GARCH and Regime switching types of models where the distribution of future outcomes change over time.

The formulation in A6 is of particular interest in our substantive context of interest because the income process literature typically assumes that log-income is additively separable in aggregate shocks (Meghir and Pistaferri, 2011). Following this literature, in this section we use log-income as the object of interest Y_i . Whenever tests require parametric assumptions on Y_{it}^B , as before we assume that beliefs about income have a step-wise uniform distribution and obtain the distribution describing beliefs about log-income accordingly.

In the income process literature, the effect of an additively separable aggregate shock is sometimes allowed to vary by education group or birth cohort. The tests described below could accommodate this if they were applied separately to each group. Note that in our empirical application we assume that the aggregate shock affects two adjacent entry cohorts of a single college in a common additive way. This does not preclude the possibility that, throughout the broader economy, the same shock differentially affects other age, education or geographic groups.

We first describe how existing mean-based tests can be modified to accommodate the possibility of aggregate shocks, and then discuss how to extend our new tests. The proofs for the Theorems that follow are in Appendix D.

6.1 Theory

Existing Tests

D'Haultfoeuille, Gaillac, and Maurel (2018) show that the DGM Test is capable of accommodating the presence of additively separable aggregate shocks. Roughly speaking, the testing procedure is as follows. First, the realization of the aggregate shock can be identified as the difference between the cross-section mean of μ_{it}^B and the cross-section mean of y_i . Then, the econometrician can test whether the distribution of $y_i + [E_i(\mu_{it}^B) - E_i(y_i)]$ is a mean-preserving spread of μ_{it}^B .

Regression Tests are based on the coefficients in the regression of realizations y_i on mean-belief μ_{it}^B . In the presence of additively separable aggregate shocks, the innovation after t ($y_i - \mu_{it}^B$) is the sum of a common aggregate shock and a residual term that is orthogonal to μ_{it}^B under Rational Expectations. Thus, y_i is the sum of the common aggregate shock, the mean-belief μ_{it}^B , and the residual term. This implies that a test can be formed based on the null hypothesis that the slope coefficient in the regression equation is equal to one.

The Correlation Test is based on the correlation between the innovation after t and

individual-specific factors that should be observed before t . By construction, aggregate shocks observed after t do not affect individual-specific factors observed before t . Moreover, as discussed above, the only effect that aggregate shocks have on the innovation is to add a common shifter to all individuals, which does not affect the correlation between the innovation and any other random variables. Hence, the Correlation Test is valid even in the presence of additively separable aggregate shocks.

Mean-Variance Test

In the presence of aggregate shocks, the difference between the cross-sectional mean of mean-beliefs, $E_i(\mu_{it}^B)$, and the cross-sectional mean of realizations, y_i , is not necessarily zero under Rational Expectations because a deviation of the aggregate shock from its mean is common to all individuals. That is, there will always exist some realization of the aggregate shock that can rationalize the observed difference in means. This implies that we cannot form a test of Rational Expectations based on the difference in means when there is a possibility of aggregate shocks after beliefs are elicited.

On the other hand, even in the presence of additively separable aggregate shocks, the Rational Expectations Hypothesis has implications for the relationship between the cross-sectional mean of subjective uncertainty, $\delta_{t,1}^B \equiv E_i[\text{var}(Y_{it}^B)]$, and the cross-sectional variance of innovations, $\delta_{t,1}^B \equiv \text{var}_i(\mu_{it}^B - y_i)$. Specifically, the former should be weakly larger than the latter. This is the case because, while the variation in the aggregate shock should be incorporated into respondents' overall subjective uncertainty, the realization of the aggregate shock is common to all individuals, and, therefore, does not contribute to the cross-sectional variance of innovations. Formally, a test of Rational Expectations can be based on the null hypothesis stated in the following theorem:

Theorem 8. *Under assumptions A1 and A6-A8, if all individuals have Rational Expectations about Y_i (Definition 1), then*

$$\delta_{t,1}^B - \delta_{t,2}^B \geq 0.$$

In contrast, as discussed in Section 4, in the absence of aggregate shocks a test can be based on the null hypothesis that $\delta_{t,1}^B = \delta_{t,2}^B$. In general, testing moment inequalities is a complicated problem because the null hypothesis includes (infinitely) many possibilities. This problem is especially difficult when we need to jointly test multiple moment inequalities. However, in our context where the test-statistic is a scalar and has a normal distribution, Karlin and Rubin (1956) states that the Uniformly Most Powerful Test for the inequality null

is the same as the one-sided test for the corresponding equality null. Hence, when aggregate shocks are potentially present, instead of basing the test on the asymptotic two-tailed p-value associated with the test-statistic \hat{u}_{12} defined in Theorem 2, we base the test on the corresponding asymptotic one-tailed p-value.

Aggregate-Distribution Test

Our discussion above suggests that, in the presence of additively separable aggregate shocks, the subjective aggregate distribution (i.e., the random variable $Y_t^B \equiv E_i(Y_{it}^B)$) and the distribution of outcomes ($Y_z \equiv Y_i | \mathbf{Z} = \mathbf{z}$) will differ even if all agents have Rational Expectations. In general, the means of the two distributions will be different because of the realization of the aggregate shock. Moreover, the subjective aggregate distribution will be more dispersed than the distribution of outcomes because the presence of aggregate shocks amplifies the mean subjective uncertainty $E_i[\text{var}(Y_{it}^B)]$, but does not affect the variance of realizations ($\text{var}_i(y_i) = \text{var}(Y_z)$). Formally, the two distributions satisfy the following relationship under Rational Expectations.

Theorem 9. *Under assumptions A1 and A6-A8, if all individuals have Rational Expectations about Y_i (Definition 1), then Y_t^B is a mean-preserving spread of $Y_z + E(Y_t^B) - E(Y_z)$.*

Theorem 9 has testable implications and can be used to form an Aggregate-Distribution Test of Rational Expectations in the presence of aggregate shocks. It is worth noting that Theorem 9 is different from the theorem that is the basis of the DGM Test. The DGM Test is based on the result that $Y_z + E(Y_t^B) - E(Y_z)$ is a mean-preserving spread of μ_{it}^B in the presence of additively separable aggregate shocks. In the absence of aggregate shocks, the null hypothesis of the Aggregate-Distribution Test encompasses the null hypothesis of the DGM test (see Section 4.4). With an additive aggregate shock this is no longer the case, and the two tests are complementary.

We first consider a nonparametric test based on Theorem 9 using information about the distribution of Y_z and the average perceived probability that y_i falls in outcome bin S_k , $E_i(P_{it}^{B,k}) = \text{Prob}(Y_t^B \in S_k)$, $k = 1, 2, \dots, K$. Given the equivalence between mean-preserving spread and second order stochastic dominance, checking whether Y_t^B is a mean-preserving spread of $Y_z + E(Y_t^B) - E(Y_z)$ involves computing the integral of the cumulative probability function of Y_t^B . However, this cumulative probability function is only partially observed because there is no restriction on how the average perceived probability $E_i(P_{it}^{B,k})$ is distributed within S_k . The partial observability of the cumulative probability function

means that a nonparametric test based on Theorem 9 requires a further restriction on that function. In particular, if 1) the union of the outcome bins spans the real line and 2) the average perceived probability $E_i(P_{it}^{B,k})$ is positive for all k , Theorem 9 does not provide any additional testable restrictions on the relationship between Y_z and $E_i(P_{it}^{B,k})$, $k = 1, 2, \dots, K$. Formally, the following theorem holds:

Theorem 10. *For any combination of Y_z , which is continuous and has a finite expectation, and $E_i(P_{it}^{B,k})$, $k = 1, 2, \dots, K$, if 1) $\bigcup_{k=1}^K S_k = \mathbb{R}$ and 2) $E_i(P_{it}^{B,k}) > 0 \forall k$, then it is always possible to find a random variable Ψ and a constant c such that:*

- (1) Ψ is a mean-preserving spread of $Y_z + c$,
- (2) $E_i(P_{it}^{B,k}) = \text{Prob}(\Psi \in S_k)$, for all k .

The intuition underlying Theorem 10 is straightforward. The assumptions imposed in Theorem 10 imply that both the highest and the lowest outcome bins are unbounded. Hence, the distribution of $E_i(P_{it}^{B,k})$ within these two outcome bins can be arbitrarily dispersed. Hence, regardless of how dispersed $Y_z + c$ is, it is always possible to find a random variable Ψ that is (1) more dispersed than $Y_z + c$ and (2) consistent with observed average perceived probabilities.

This intuition also suggests that it is possible to reject Rational Expectations based on Theorem 9 if the union of outcome bins, $\bigcup_{k=1}^K S_k = \mathbb{R}$, is bounded. In this case, if a random variable Ψ is consistent with $E_i(P_{it}^{B,k})$, $k = 1, 2, \dots, K$, its variance is bounded above. If this upper bound is smaller than the variance of Y_z , then Ψ cannot be a mean-preserving spread of $Y_z + c$, which indicates a rejection of Rational Expectations. We provide one such example in Appendix F.

Theorem 10 suggests that, when eliciting expectations using survey questions, it would be desirable to ensure that the union of outcome bins is bounded. This might be achieved by combining the probability format with respondent-reported (finite) lower and upper bounds.

In our context, the outcome of interest is log-income. As seen in Survey Question A2 in Appendix A, the lowest outcome bin, $(-\infty, \log(15,000))$, is not bounded below. While the highest outcome bin, $(\log(135,000), \log(1,000,000))$, is technically bounded above, it is not clear whether respondents treat this upper bound as a real restriction. For example, respondents were instructed that the probabilities that they allocate to the outcome bins should sum up to one. Hence, if a person thought there was some chance that her family would earn more than \$1,000,000, she would likely add the probability that this happens to

the highest outcome bin provided by the survey question. Further, we find that the average perceived probabilities are positive for all outcome bins. Then, Theorem 10 implies that it is not feasible to form an informative nonparametric test of Rational Expectations based on Theorem 9 in our case. Therefore, we turn to a parametric version of the Aggregate-Distribution test.

As mentioned in Section 4.3.2, the mean component of the parametric Aggregate-Distribution Test is essentially the same as the test that there is no difference between the cross-sectional mean of mean-beliefs and the cross-sectional mean of realizations (no mean bias). Therefore, as is the case with the mean component of the Mean-Variance test, it is not possible to form a test of Rational Expectations by modifying the mean component of parametric Aggregate-Distribution Test.

However, consistent with our discussion above for the Mean-Variance Test, a test of Rational Expectations can be formed by modifying the variance component of the Parametric Aggregate-Distribution Test. Theorem 9 implies that the variance of the subjective aggregate distribution, $E_i(\text{var}(Y_{it}^B)) + \text{var}_i(\mu_{it}^B)$, is larger than the variance of outcomes, $\text{var}_i(y_i)$, if additively separable aggregate shocks are present. Thus, similar to the modified Mean-Variance Test, we base the test on the asymptotic one-tailed p-value associated with the test-statistic \hat{u}_V .

6.2 Empirical Results

We now apply these modified tests to both in-school and post-college beliefs. In contrast to Section 5, we now work with beliefs about, and realizations of, the logarithm of annual family income. So the tests allow common additive shocks to log-income, consistent with much of the income process literature. Again we note that aggregate shocks are assumed to affect two adjacent cohorts of students from a single college in a common way. This does not preclude differential effects on other age, education or geographic groups.

The results are presented in Table 8, which reports test of the rationality of beliefs about log-income at age 28 or 29 at different time horizons that cover both the in-school and post-college periods. Throughout, we restrict attention to the balanced panel.

Starting with Table 8, for the variance component of the Mean-Variance test (row 1), at each of these in-school points, we reject the null hypothesis that $\delta_{t,1}^B - \delta_{t,2}^B \geq 0$ at a 1% level. In contrast, in the post-college period, we cannot reject this null hypothesis at a 5% level for either 1-year ahead or 5-year ahead beliefs about log-income at age 28/29. In some

instances, the statistics are positive and quite large, leading to p-values close to 1.

The Correlation Test (row 5), and the variance component of the Parametric Aggregate-Distribution Test (row 2) produce results that are broadly similar, almost always rejecting for the in-school period, and not rejecting in the post-college period. For example, for the Parametric Aggregate-Distribution Test, while we cannot reject the null hypothesis that $E_i(\text{var}(Y_{it}^B)) + \text{var}_i(\mu_{it}^B) - \text{var}_i(y_i) \geq 0$ at a 5% level for beliefs elicited at the time of entrance, this null hypothesis is rejected at a 5% level for beliefs elicited at the end of each of the four academic years.¹³ However, we cannot reject the null hypothesis at a 5% level for 1-year ahead beliefs nor 5-year ahead income beliefs in the post college period.

The DGM Test (row 3) and Regression test (row 4) produce different patterns. In contrast to the results in Section 5.1, the DGM Test does not reject Rational Expectations at a 5% level for any of the in-school beliefs once we allow for additively separable aggregate shocks (and change the object of interest to log-income). It now fails to reject Rational Expectations both in-school and post-college. In contrast, the Regression Test rejects at the 1% for every time horizon, both in-school and post-college.¹⁴

Broadly, the conclusion from Section 5.1 holds when we allow for aggregate shocks. The evidence against the Rational Expectations Hypothesis is much stronger for the in-school period than for the post-college period. Interestingly, after allowing for aggregate shocks, the DGM test tends not to reject Rational Expectations in our data.¹⁵ The only rejections we find for the post-college period are for the Regression Test. However, we note again that measurement error in μ_{it}^B , arising either from reporting error in beliefs or approximation error in computing individual means, can lead to spurious rejections with this test. On the other hand, for the post-college period, the Parametric Aggregate-Distribution test is based on beliefs elicited in the probability format, and so is robust to classical reporting errors as discussed in Section 4.3.2. With this test, we cannot reject Rational Expectations in the post-college period.

¹³We can reject the null hypothesis at a 10% level for beliefs elicited at the time of entrance.

¹⁴Note that in the presence of aggregate shocks, Theorem 7 does not apply, so the null of the regression test is not implied by combination of the nulls of the Mean-Variance and Aggregate-Distribution tests.

¹⁵We note again that our implementation of the test imputes a mean from probabilistic expectations data, rather than interpreting a reported point belief as a mean.

Table 7: Parametric Tests about Income in Year 2007-2013 (p-values)

Year of Realization	2007	2008	2009	2010	2011	2012	2013
Panel A: 1-year ahead Beliefs							
\hat{u}_{12} (uncertainty)	0.0730	0.0055	0.0059	0.0459	0.0303	0.0009	0.0028
\hat{u}_3 (mean)	0.4448	0.2299	0.3736	0.4750	0.3902	0.0000	0.0056
\hat{u}_V (variance)	0.0137	0.3669	0.5582	0.2496	0.6849	0.1129	0.0060
Joint	0.0007	0.0223	0.0364	0.2676	0.1926	$< 10^{-4}$	0.0114
Sample Size	200	264	269	294	320	322	344
Panel B: 5-year ahead Beliefs							
\hat{u}_{12} (uncertainty)	N.A.	N.A.	N.A.	N.A.	$< 10^{-4}$	$< 10^{-4}$	0.0002
\hat{u}_3 (mean)	N.A.	N.A.	N.A.	N.A.	0.0134	0.2218	0.7128
\hat{u}_V (variance)	N.A.	N.A.	N.A.	N.A.	0.8124	0.2787	0.0198
Joint	N.A.	N.A.	N.A.	N.A.	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
Sample Size	N.A.	N.A.	N.A.	N.A.	184	237	265

Note: Panel A reports the p-values associated with (two-tailed) test statistics \hat{u}_{12} , \hat{u}_3 (defined in Theorem 2), \hat{u}_V (defined in Theorem 6), and the joint test in Section 4.4 for students' 1-year ahead beliefs after college about annual family income realized in 2007-2013. Panel B reports the p-values associated with (two-tailed) test statistics \hat{u}_{12} , \hat{u}_3 , \hat{u}_V , and the joint test for students' 5-year ahead beliefs after college about annual family income realized in 2011-2013.

Table 8: Rational Expectations Tests about Income at Age 28-29 with Aggregate Shocks

Year of Belief-elicitation	10-year ahead	9-year ahead	8-year ahead	7-year ahead	6-year ahead	5-year ahead	1-year ahead
Mean-Variance Test p-value	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	0.1461	0.8635
Aggregate-Distribution Test p-value	0.0978	0.0138	0.0014	0.0125	0.0235	0.9998	0.9940
DGM Test p-value	1.0000	1.0000	1.0000	1.0000	1.0000	0.1543	0.5952
Regression Test: Slope	0.2200	0.4132	0.3362	0.3837	0.5691	0.5855	0.8913
Regression Test: p-value	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	0.0093
Correlation Test: Correlation	N.A.	0.2320	0.1924	0.2482	0.1767	-0.0784	0.0262
Correlation Test: p-value	N.A.	0.0001	0.0062	0.0004	0.0100	0.3435	0.6545

Notes: All tests are applied to balanced panels. The first five columns report the results for aggregate-shock-robust tests described in Section 6.1 for students' beliefs in college about own annual income at age 28. The last two columns report the results for aggregate-shock-robust tests for students' beliefs after college about annual family income at age 28 (or 29).

7 Conclusion

The key point of this paper is that quality of individuals' beliefs depends on the higher moments of those beliefs, not just the means. Ex post accuracy (in a belief-weighted mean-squared error sense) depends on uncertainty as well as bias. Similarly, ex ante rationality implies not only that a person's beliefs are unbiased but that the entire distribution describing a person's subjective beliefs about a future variable should equal the actual conditional distribution of that variable, given the person's information set. Particularly with a continuous variable such as income, this requirement can be violated even if beliefs are unbiased. An individual may have correct mean-beliefs about future income but underestimate the income risk they face. This would violate Rational Expectations, and could lead to sub-optimal choices. Such an individual may be insufficiently prudent in their consumption-savings choices, for example.

It is becoming more common to collect information in surveys about the distributions describing individuals' beliefs. We propose a new micro-data measure of ex post accuracy, and new tests of Rational Expectations, designed to exploit information about the distribution describing a person's beliefs. Our new tests complement tests that are based on only the means of belief distributions, such as the well-known regression test, or the new test of D'Haultfoeuille, Gaillac, and Maurel (2018). The latter takes account of all of the information in the distribution of mean beliefs (across persons) and the distribution of outcomes. Our new tests exploit the additional information in each person's belief distribution.

We apply this measure and these tests to data on income expectations and realizations from the BPS. The BPS is unique in allowing a comparison of probabilistic income expectations and income realizations over an extended age range: from college entrance through early experiences in the labor market. We find that individuals' beliefs about future income become more ex post accurate through college and into the post-college period both because the average prediction error falls, and because uncertainty resolves. Applied to the BPS data our new tests reject ex ante Rational Expectations for college students, even when we allow for additive aggregate shocks. Importantly, we find that our new tests reject in cases where existing mean-based tests do not. In contrast to our findings for college students, we find much less evidence against Rational Expectations once these students have graduated, and most have had their first experience in the labor market.

Our analysis offers a number of lessons for survey designers. First, this work reinforces the value of collecting information about belief distributions, and not just point expectations.

Second, our analysis highlights a benefit of collecting such data in the probability format, and with common cutoffs across respondents: this design facilitates the construction of a nonparametric distribution test. Randomized or respondent-driven cutoffs are sometimes advocated as a way to avoid anchoring effects (e.g., Manski, 2004), but a nonparametric distribution test would then involve many moment inequalities, so there is a trade-off to be considered. Finally, we demonstrated that in the presence of (additive) aggregate shocks, a nonparametric aggregate distribution test requires informative upper and lower bounds on the belief distribution. That is something that can be built into survey questions measuring beliefs.

The analyses in this paper could be extended in a number of ways. Benítez-Silva et al. (2008) construct mean-based tests of Rational Expectations from observed innovations in beliefs (rather than from pairings of beliefs and realizations.) An aggregate distribution test could be constructed similarly, from a sequence of observations on beliefs.

In a recent paper, Hahn, Kuersteiner, and Mazzocco (2019) study the estimation of the parameters of economic models with Rational Expectations, in environments with aggregate shocks. They show that model parameters (such as taste or technology parameters) can not be consistently estimated with cross-sectional (or short panel) data simply by conditioning on the realization of the aggregate shock. The aggregate distribution test we propose for an environment with aggregate shocks does not contradict this result. Rational Expectations can still place restrictions on the data, particularly inequality restrictions, even if those restrictions are not sufficient to identify model parameters. Nevertheless, the methods developed in Hahn, Kuersteiner, and Mazzocco (2019) might be used to combine cross-sectional information on probabilistic expectations and individual realizations with a long time series of realizations of the aggregate shock. This could generate additional testable implications of Rational Expectations and more powerful tests.

Using data from the Health and Retirement Study, Giustinelli, Manski, and Molinari (2019) show that individuals report intervals for perceived probabilities of future events when allowed to do so. They interpret this finding as evidence that individuals have imprecise beliefs, also known as ambiguity, about future outcomes. We note that having imprecise beliefs is itself a deviation from Rational Expectations. Because the BPS was not designed to elicit interval subjective probabilities, we are unable to directly assess the precision of respondents' income beliefs. Nonetheless, our Rational Expectations tests, which are based on point probabilities (for each interval of the range of outcomes) and/or percentiles, can have

meaningful interpretations under additional assumptions about these point probabilities and percentiles.¹⁶ For example, while an individual might not be able to precisely choose among multiple possible subjective distributions, she may believe that one of these distributions is more likely to be true than others. Then, if everyone reports probabilities/percentiles associated with this most likely subjective distribution, our tests in effect examine whether the most likely subjective income distributions are consistent with the implications of Rational Expectations.

We leave further exploration of these possible extensions for future work.

Funding

This project was made possible by generous support from the Mellon Foundation, the Spencer Foundation, the National Science Foundation, and the Social Sciences and Humanities Research Council.

Conflicts of Interest

None.

References

- Ameriks, John, Gábor Kézdi, Minjoon Lee, and Matthew D Shapiro. 2020. “Heterogeneity in expectations, risk tolerance, and household stock shares: The attenuation puzzle.” *Journal of Business & Economic Statistics* 38 (3):633–646.
- Arcidiacono, Peter. 2004. “Ability sorting and the returns to college major.” *Journal of Econometrics* 121 (1-2):343–375.
- Arcidiacono, Peter, V Joseph Hotz, and Songman Kang. 2012. “Modeling college major choices using elicited measures of expectations and counterfactuals.” *Journal of Econometrics* 166 (1):3–16.

¹⁶Note the distinction between *point probabilities* for each interval in a partition of the range of outcomes, and the single *point expectation* that is sometimes interpreted as the mean of the distribution describing beliefs.

- Attanasio, Orazio and Britta Augsburg. 2016. "Subjective expectations and income processes in rural India." *Economica* 83 (331):416–442.
- Benítez-Silva, Hugo, Debra S Dwyer, Wayne-Roy Gayle, and Thomas J Muench. 2008. "Expectations in micro data: rationality revisited." *Empirical Economics* 34 (2):381–416.
- Bowen, William G and Derek Bok. 1998. *The Shape of the River. Long-Term Consequences of Considering Race in College and University Admissions*. ERIC.
- Bowen, William G, Matthew M Chingos, and Michael S McPherson. 2009. *Crossing the finish line: Completing college at America's public universities*. Princeton University Press.
- Das, Marcel, Jeff Dominitz, and Arthur Van Soest. 1999. "Comparing predictions and outcomes: Theory and application to income changes." *Journal of the American Statistical Association* 94 (445):75–85.
- Das, Marcel and Bas Donkers. 1999. "How certain are Dutch households about future income? An empirical analysis." *Review of Income and Wealth* 45 (3):325–338.
- Das, Marcel and Arthur Van Soest. 1997. "Expected and realized income changes: Evidence from the Dutch socio-economic panel." *Journal of Economic Behavior & Organization* 32 (1):137–154.
- Delavande, Adeline. 2008. "Pill, patch, or shot? Subjective expectations and birth control choice." *International Economic Review* 49 (3):999–1042.
- D'Haultfoeuille, Xavier, Christophe Gaillac, and Arnaud Maurel. 2018. "Rationalizing Rational Expectations? Tests and Deviations." (25274).
- Dominitz, J and CF Manski. 1997a. "Perceptions of economic insecurity: Evidence from the survey of economic expectations." *Public opinion quarterly* 61 (2):261–287.
- Dominitz, Jeff. 1998. "Earnings Expectations, Revisions, and Realizations." *The Review of Economics and Statistics* 80 (3):374–388.
- Dominitz, Jeff and Charles F Manski. 1996. "Eliciting student expectations of the returns to schooling." *The Journal of Human Resources* 31 (1):1.
- Dominitz, Jeff and Charles F. Manski. 1997b. "Using Expectations Data to Study Subjective Income Expectations." *Journal of the American Statistical Association* 92 (439):855–867.

- Giustinelli, Pamela, Charles F Manski, and Francesca Molinari. 2019. “Precise or Imprecise Probabilities? Evidence from Survey Response on Late-onset Dementia.” (26125). URL <http://www.nber.org/papers/w26125>.
- Gong, Yifan, Ralph Stinebrickner, and Todd Stinebrickner. forthcoming. “Marriage, children, and labor supply: Beliefs and outcomes.” *Journal of Econometrics* .
- Gong, Yifan, Todd Stinebrickner, and Ralph Stinebrickner. 2019. “Uncertainty about future income: Initial beliefs and resolution during college.” *Quantitative Economics* 10 (2):607–641.
- . 2020. “Perceived and actual option values of college enrollment.” *Journal of Applied Econometrics* 35 (7):940–959.
- Guiso, Luigi, Tullio Jappelli, and Daniele Terlizzese. 1992. “Earnings uncertainty and precautionary saving.” *Journal of Monetary Economics* 30 (2):307–337.
- Hahn, Jinyong, Guido Kuersteiner, and Maurizio Mazzocco. 2019. “Estimation with Aggregate Shocks.” *The Review of Economic Studies* 87 (3):1365–1398.
- Karlin, Samuel and Herman Rubin. 1956. “The theory of decision procedures for distributions with monotone likelihood ratio.” *The Annals of Mathematical Statistics* :272–299.
- Leland, Hayne E. 1978. “Saving and uncertainty: The precautionary demand for saving.” In *Uncertainty in economics*. Elsevier, 127–139.
- Manski, Charles F. 1992. *Parental income and college opportunity*. Democratic Study Center.
- . 2004. “Measuring expectations.” *Econometrica* 72 (5):1329–1376.
- . 2018. “Survey measurement of probabilistic macroeconomic expectations: progress and promise.” *NBER Macroeconomics Annual* 32 (1):411–471.
- Manski, Charles F and Francesca Molinari. 2010. “Rounding probabilistic expectations in surveys.” *Journal of Business & Economic Statistics* 28 (2):219–231.
- Meghir, Costas and Luigi Pistaferri. 2011. “Earnings, consumption and life cycle choices.” In *Handbook of labor economics*, vol. 4. Elsevier, 773–854.

- Pesaran, Hashem and Martin Weale. 2006. "Survey expectations." *Handbook of economic forecasting* 1:715–776.
- Sandmo, Agnar. 1970. "The effect of uncertainty on saving decisions." *The review of economic studies* 37 (3):353–360.
- Stinebrickner, Ralph and Todd Stinebrickner. 2008. "The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study." *American Economic Review* 98 (5):2163–84.
- . 2014a. "Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model." *Journal of Labor Economics* 32 (3):601–644.
- Stinebrickner, Ralph, Todd Stinebrickner, and Paul Sullivan. 2019. "Job tasks, time allocation, and wages." *Journal of Labor Economics* 37 (2):399–433.
- Stinebrickner, Ralph and Todd R Stinebrickner. 2003. "Understanding educational outcomes of students from low-income families evidence from a liberal arts college with a full tuition subsidy program." *Journal of Human Resources* 38 (3):591–617.
- . 2014b. "A major in science? Initial beliefs and final outcomes for college major and dropout." *Review of Economic Studies* 81 (1):426–472.
- Stinebrickner, Todd and Ralph Stinebrickner. 2012. "Learning about academic ability and the college dropout decision." *Journal of Labor Economics* 30 (4):707–748.
- Van der Klaauw, Wilbert. 2012. "On the use of expectations data in estimating structural dynamic choice models." *Journal of Labor Economics* 30 (3):521–554.
- Zafar, Basit. 2011. "How do college students form expectations?" *Journal of Labor Economics* 29 (2):301–348.

Appendices

A Survey Questions

Question A1. The following questions will ask you about the income you might earn in the future at different ages under several hypothetical scenarios. We realize that you will not know exactly how much money you would make at a particular point in time. However, you may believe that some amounts of money are quite likely while others are quite unlikely. We would like to know what you think. 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. When reporting incomes, take into account the possibility that you will work full-time, the possibility that you will work part-time, the possibility that you will not be working, and (for the hypothetical scenarios which involve graduation) the possibility that you will attend graduate or professional school. When reporting income you should ignore the effects of price inflation.

For ALL of question 1A, **assume that you graduate from Berea.** Think about the kinds of jobs that will be available for you and those that you would accept. 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

Question A2. The total yearly earnings of your family **as of today** are shown in Question A3. We know that you do not know exactly what the total yearly earnings for your family will be **one year from today** because you or your spouse (if you have one) may get raises on your current jobs, take pay cuts on your current jobs, lose your current jobs, change jobs, quite your jobs, or take a leave from your current jobs. Taking into account uncertainty about these possibilities, we want to know what you believe about your earnings one year from now. 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**. Please take into account that your jobs situations may change in the next year and also take into account changes in the number of workers in your family due to marriage or the birth of new children. **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.

<u>Earnings Interval</u>	<u>Percent Chance</u>
[\$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)	_____

NOTE: Similar questions were asked about total yearly family income **five years from today**.

Question A3. 1) How many hours do you typically work each week in your job? (a)
 2) Approximately how much do you earn in your job? NOTE: Please indicate both a dollar amount and whether this amount is your pay per hour, per day, per week, per month, per year etc. For example, if you earn \$8.50 an hour, please write \$8.50 per hour. If you earn \$30,000 per year, please write \$30,000 per year. \$ (b) per (c)

We use the numbers in (a), (b), and (c) to compute your yearly earnings. The formula

you should use depends on whether you reported your earnings per Year, per MONTH, per WEEK, per HOUR etc. in (c).

If per YEAR, then your YEARLY EARNINGS is the \$ amount from (b): b

If per MONTH, then your YEARLY EARNINGS is the \$ amount from (b) multiplied by 12: $b \times 12$

If per TWO WEEKS, then your YEARLY EARNINGS is the \$ amount from (b) multiplied by 26: $b \times 26$

If per WEEK, then your YEARLY EARNINGS is the \$ amount from (b) multiplied by 52: $b \times 52$

If per HOUR, then your YEARLY EARNINGS is the \$ amount from (b) multiplied by 52 and then multiplied by the HOURS number in (a): $b \times 52 \times a$

NOTE: Similar questions were asked about total yearly **spousal** income.

Question A4. We are interested in whether you think you will get married and when you think you will get married. What is the percent chance that your first marriage will take place at each of the following ages or not at all? **Note: Each number should be between 0 and 100 and the numbers should sum to 100.**

Your Age	Percent Chance of first marriage taking place at this age
At or before Age 23	_____
At Age 24 or 25	_____
At Age 26 or 27	_____
At Age 28 or 29	_____
At or after Age 30	_____
Never get married	_____

B Stepwise Uniform Distribution

Regardless of the format we use to elicit probabilistic expectations, for each respondent, we obtain a set of outcome intervals and corresponding perceived probabilities $\{(S_k, P_{it}^{B,k}) | k = 1, 2, \dots, K\}$. If the beliefs are elicited in a percentile format, then $P_{it}^{B,k}$ are pre-defined by researchers and S_k are reported by the respondent. Alternatively, if the beliefs are elicited in a probability format, then S_k are pre-defined by researchers and $P_{it}^{B,k}$ are reported by the respondent. We let $C_{k,l}$ and $C_{k,u}$ denote the infimum and the supremum of interval S_k .

We assume that the respondent's subjective beliefs Y_{it}^B has the stepwise uniform pdf given by:

$$f_{Y_{it}^B}(y_{it}^B) = \frac{P_{it}^{B,k}}{C_{k,u} - C_{k,l}}, \text{ if } y_{it}^B \in S_k, \text{ for } k \in \{1, 2, \dots, K\}. \quad (15)$$

The mean and variance of Y_{it}^B are given by $E(Y_{it}^B) = \sum_{k=1}^K \frac{P_{it}^{B,k}(C_{k,u} - C_{k,l})}{2}$ and $var(Y_{it}^B) = \sum_{k=1}^K \frac{P_{it}^{B,k}[(C_{k,u})^2 + C_{k,u}C_{k,l} + (C_{k,l})^2]}{3} - (E(Y_{it}^B))^2$.

C Converting Beliefs about Own Income to Beliefs about Family Income

We begin by assuming that $Y_{O,it}^B$ and $Y_{S,it}^B$ are identically distributed. In our framework, this would follow if mating were (perfectly) assortative on factors that are observed at t . We further initially assume that $Y_{O,it}^B$, $Y_{S,it}^B$, and P_{it}^M are mutually independent. Under these assumptions, we can show that $E(Y_{F,it}^B)$ and $var(Y_{F,it}^B)$ are functions of objects, P_{it}^M , $E(Y_{O,it}^B)$ and $var(Y_{O,it}^B)$, which are directly available in the BPS,

$$E(Y_{F,it}^B) = (1 + P_{it}^M)E(Y_{O,it}^B), \quad (16)$$

$$var(Y_{F,it}^B) = (1 + P_{it}^M)var(Y_{O,it}^B) + [E(Y_{O,it}^B)]^2 P_{it}^M (1 - P_{it}^M). \quad (17)$$

If students believe that mating will also positively sort on income-influencing factors that are still unresolved at time t , then $Y_{O,it}^B$ and $Y_{S,it}^B$ will be positively correlated. In this case, Equation (16) correctly states $E(Y_{F,it}^B)$, while Equation (17) understates $var(Y_{F,it}^B)$. This is of no consequence for $\delta_{F,t,2}^B$ and $\delta_{F,t,3}^B$, but suggests our calculation of $\delta_{F,t,1}^B$ (and therefore $\Delta_{F,t}^B$) should be viewed as a *lower* bound.

D Proofs

Theorem 1. *Under assumptions A1 to A3, if all individuals have Rational Expectations about Y_i (Definition 1), then*

$$\delta_{t,1}^B = \delta_{t,2}^B \text{ and } \sqrt{\delta_{t,3}^B} = 0.$$

Proof. Consider an individual i of type θ_i . Under the assumption that she has Rational

Expectations about Y_i , her beliefs at t about y_i , Y_{it}^B , are given by:

$$Y_{it}^B = Y_i | (\mathbf{X}_i^{\theta_i, t^-} = \mathbf{x}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}). \quad (18)$$

Under the assumption that Θ_i and \mathbf{X}_i are independent (A3), $\Theta_i = \theta_i$ does not provide additional information about Y_i , conditional on $\mathbf{X}_i^{\theta_i, t^-}$ and \mathbf{Z}^{t^-} . Hence, the following equation holds:

$$Y_{it}^B = Y_i | (\Theta_i = \theta_i, \mathbf{X}_i^{\theta_i, t^-} = \mathbf{x}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}). \quad (19)$$

This implies that:

$$\mu_{it}^B = E(Y_{it}^B) = E[Y_i | (\Theta_i = \theta_i, \mathbf{X}_i^{\theta_i, t^-} = \mathbf{x}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-} = \mathbf{z}^{t^-})], \quad (20)$$

$$\text{var}(Y_{it}^B) = \text{var}[Y_i | (\Theta_i = \theta_i, \mathbf{X}_i^{\theta_i, t^-} = \mathbf{x}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-} = \mathbf{z}^{t^-})]. \quad (21)$$

The population distribution of Y_{it}^B is produced by the realizations of Θ_i and $\mathbf{X}_i^{\theta_i, t^-}$. Hence, the first component, $\delta_{t,1}^B = E_i[\text{var}(Y_{it}^B)]$, is given by:

$$\delta_{t,1}^B = E\{\text{var}[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-})] | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}\}. \quad (22)$$

Similarly, the population distribution of y_i is produced by the realizations of \mathbf{X}_i . Hence, the second component, $\delta_{t,2}^B = \text{var}_i(y_i - \mu_{it}^B)$, is given by:

$$\begin{aligned} \delta_{t,2}^B &= \text{var}\{Y_i - E[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-})] | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}\} \\ &= \text{var}\{E\{Y_i - E[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-})] | (\Theta_i, \mathbf{Z}^{t^-})\} | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}\} + \\ &\quad E\{\text{var}\{Y_i - E[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-})] | (\Theta_i, \mathbf{Z}^{t^-})\} | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}\} \\ &= \text{var}(0 | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}) + E\{\text{var}\{Y_i - E[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-})] | (\Theta_i, \mathbf{Z}^{t^-})\} | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}\} \\ &= E\{\text{var}\{E\{Y_i - E[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-})] | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-})\} | (\Theta_i, \mathbf{Z}^{t^-})\} | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}\} + \\ &\quad E\{E\{\text{var}\{Y_i - E[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-})] | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-})\} | (\Theta_i, \mathbf{Z}^{t^-})\} | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}\} \\ &= E\{\text{var}[0 | (\Theta_i, \mathbf{Z}^{t^-})] | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}\} + E\{E\{\text{var}[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-})] | (\Theta_i, \mathbf{Z}^{t^-})\} | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}\} \\ &= E\{\text{var}[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-})] | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}\} \\ &= \delta_{t,1}^B, \end{aligned} \quad (23)$$

where the second and fourth lines follow from the law of iterative variance.

The second part of Theorem 1 directly follows from the law of iterative expectations:

$$\sqrt{\delta_{t,3}^B} = \bar{y} - \bar{\mu}_t = E(Y_i | \mathbf{Z}^{t-} = \mathbf{z}^{t-}) - E\{E[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t-}, \mathbf{Z}^{t-})] | \mathbf{Z}^{t-} = \mathbf{z}^{t-}\} = 0. \quad (24)$$

□

Theorem 2. (*Mean-Variance Test*) Under assumptions A1 to A4, if all individuals have Rational Expectations about Y_i (Definition 1), then

$$\hat{u}_{12} \equiv \frac{\hat{\delta}_{t,1}^B - \hat{\delta}_{t,2}^B}{\hat{\sigma}_{12}/\sqrt{n}} \xrightarrow{d} N(0, 1),$$

$$\hat{u}_3 \equiv \frac{\sqrt{\hat{\delta}_{t,3}^B}}{\hat{\sigma}_3/\sqrt{n}} \xrightarrow{d} N(0, 1).$$

Proof. Theorem 2 trivially follows from the standard central limit theorem. □

Theorem 3. Under assumptions A1 to A3, if all individuals have Rational Expectations about Y_i (Definition 1), then

$$E_i(P_{it}^{B,k}) = \text{Prob}(Y_i \in S_k | \mathbf{Z}^{t-} = \mathbf{z}^{t-}), \quad \text{for all } k.$$

Proof. Let $I(\cdot)$ denote an indicator function that takes a value of one when the statement is true and takes a value of zero when the statement is false. Consider an individual i of type θ_i . Under the assumption that she has Rational Expectations about Y_i , her perceived probability $P_{it}^{B,k}$ is given by:

$$\begin{aligned} P_{it}^{B,k} &= \text{Prob}(Y_i \in S_k | (\mathbf{X}_i^{\theta_i, t-} = \mathbf{x}_i^{\theta_i, t-}, \mathbf{Z}^{t-} = \mathbf{z}^{t-})) \\ &= E[I(Y_i \in S_k) | (\mathbf{X}_i^{\theta_i, t-} = \mathbf{x}_i^{\theta_i, t-}, \mathbf{Z}^{t-} = \mathbf{z}^{t-})]. \end{aligned} \quad (25)$$

Similar as before, the independence between Θ_i and \mathbf{X}_i implies that:

$$\begin{aligned} P_{it}^{B,k} &= \text{Prob}(Y_i \in S_k | (\Theta_i = \theta_i, \mathbf{X}_i^{\theta_i, t-} = \mathbf{x}_i^{\theta_i, t-}, \mathbf{Z}^{t-} = \mathbf{z}^{t-})) \\ &= E[I(Y_i \in S_k) | (\Theta_i = \theta_i, \mathbf{X}_i^{\theta_i, t-} = \mathbf{x}_i^{\theta_i, t-}, \mathbf{Z}^{t-} = \mathbf{z}^{t-})]. \end{aligned} \quad (26)$$

The population distribution of $P_{it}^{B,k}$ is produced by the realizations of Θ_i and $\mathbf{X}_i^{\theta_i, t-}$.

Hence, $E_i(P_{it}^{B,k})$ is given by:

$$\begin{aligned}
E_i(P_{it}^{B,k}) &= E\{E[I(Y_i \in S_k)|(\Theta_i, \mathbf{X}_i^{\theta_i, t-}, \mathbf{Z}^{t-})|\mathbf{Z}^{t-} = \mathbf{z}^{t-}]\} \\
&= E[I(Y_i \in S_k)|\mathbf{Z}^{t-} = \mathbf{z}^{t-}] \\
&= \text{Prob}(Y_i \in S_k|\mathbf{Z}^{t-} = \mathbf{z}^{t-}),
\end{aligned} \tag{27}$$

where the second line directly follows from the law of iterative expectations. \square

Theorem 4. (Nonparametric Aggregate-Distribution Test) Denote the mean and variance of $\tilde{\mathbf{P}}_{it}^B$ for sample $\{\mathcal{P}_I, \mathcal{P}_{II}\}$ as $\hat{\mu}_P$ and \hat{V}_P , respectively, and denote the mean and variance of \mathbf{Q}_i for sample $\{\mathcal{Q}_I, \mathcal{Q}_{II}\}$ as $\hat{\mu}_Q$ and \hat{V}_Q , respectively. Let \hat{V}_{PQ} and \hat{V}_{QP} denote the cross-covariance matrices of $\tilde{\mathbf{P}}_{it}^B$ and \mathbf{Q}_i for sample $\{(\tilde{\mathbf{P}}_{it}^B, \mathbf{Q}_i)|\tilde{\mathbf{P}}_{it}^B \in \mathcal{P}_I, \mathbf{Q}_i \in \mathcal{Q}_I\}$. Finally, we denote the number of individuals who reported both probabilities and realizations as n_I^{PQ} , and denote the number of individuals who are in \mathcal{P}_{II} and \mathcal{Q}_{II} , respectively, as n_{II}^P and n_{II}^Q , respectively.

Under assumptions A1 to A3, if all individuals have Rational Expectations about Y_i (Definition 1), then

$$\hat{u}_p \equiv (\hat{\mu}_P - \hat{\mu}_Q)' \hat{V}^{-1} (\hat{\mu}_P - \hat{\mu}_Q) \xrightarrow{d} \chi^2(K-1),$$

$$\text{where } \hat{V} = \frac{\hat{V}_P}{n_I^{PQ} + n_{II}^P} + \frac{\hat{V}_Q}{n_I^{PQ} + n_{II}^Q} - \frac{n_I^{PQ}}{(n_I^{PQ} + n_{II}^P)(n_I^{PQ} + n_{II}^Q)} (\hat{V}_{PQ} + \hat{V}_{QP}).$$

Proof. Denote the population mean and variance of $\tilde{\mathbf{P}}_{it}^B$ as μ_P and V_P , respectively, and denote the population mean and variance of \mathbf{Q}_i as μ_Q and V_Q , respectively. Let V_{PQ} and V_{QP} denote the population cross-covariance matrices of $\tilde{\mathbf{P}}_{it}^B$ and \mathbf{Q}_i .

We rewrite $\hat{\mu}_P$ as $\frac{n_I^{PQ}}{n_I^{PQ} + n_{II}^P} \hat{\mu}_P^I + \frac{n_{II}^P}{n_I^{PQ} + n_{II}^P} \hat{\mu}_P^{II}$, where $\hat{\mu}_P^I$ and $\hat{\mu}_P^{II}$ are the averages of $\tilde{\mathbf{P}}_{it}^B$ for sample \mathcal{P}_I and \mathcal{P}_{II} , respectively. Similarly, we rewrite $\hat{\mu}_Q$ as $\frac{n_I^{PQ}}{n_I^{PQ} + n_{II}^Q} \hat{\mu}_Q^I + \frac{n_{II}^Q}{n_I^{PQ} + n_{II}^Q} \hat{\mu}_Q^{II}$, where $\hat{\mu}_Q^I$ and $\hat{\mu}_Q^{II}$ are the averages of \mathbf{Q}_i for sample \mathcal{Q}_I and \mathcal{Q}_{II} , respectively.

Under the null of Rational Expectations, we have $\mu_P = \mu_Q$. Hence, we can write $\hat{\mu}_P - \hat{\mu}_Q$ as a linear combination of $\hat{\mu}_P^I - \mu_P$, $\hat{\mu}_P^{II} - \mu_P$, $\hat{\mu}_Q^I - \mu_Q$, and $\hat{\mu}_Q^{II} - \mu_Q$ as follows.

$$\hat{\mu}_P - \hat{\mu}_Q = \left(\frac{n_I^{PQ}}{n_I^{PQ} + n_{II}^P}, \frac{n_{II}^P}{n_I^{PQ} + n_{II}^P}, -\frac{n_I^{PQ}}{n_I^{PQ} + n_{II}^Q}, -\frac{n_{II}^Q}{n_I^{PQ} + n_{II}^Q} \right) \begin{bmatrix} \hat{\mu}_P^I - \mu_P \\ \hat{\mu}_P^{II} - \mu_P \\ \hat{\mu}_Q^I - \mu_Q \\ \hat{\mu}_Q^{II} - \mu_Q \end{bmatrix}. \tag{28}$$

Under the assumption of Missing Completely at Random, $(\frac{n_I^{PQ}}{n_I^{PQ}+n_{II}^P}, \frac{n_{II}^P}{n_I^{PQ}+n_{II}^P}, -\frac{n_I^{PQ}}{n_I^{PQ}+n_{II}^Q}, -\frac{n_{II}^Q}{n_I^{PQ}+n_{II}^Q})$ converges in probability to a certain vector, which we denote as \mathbf{R} .

Under the MCAR assumption, other than the combination of $\hat{\mu}_P^I - \mu_P$ and $\hat{\mu}_Q^I - \mu_Q$, the cross-covariance matrix between any two of $\hat{\mu}_P^I - \mu_P$, $\hat{\mu}_P^{II} - \mu_P$, $\hat{\mu}_Q^I - \mu_Q$, and $\hat{\mu}_Q^{II} - \mu_Q$ is zero. Hence, the central limit theorem implies that the asymptotic distribution of the vector $(\hat{\mu}_P^I - \mu_P, \hat{\mu}_P^{II} - \mu_P, \hat{\mu}_Q^I - \mu_Q, \hat{\mu}_Q^{II} - \mu_Q)$ is given by:

$$\begin{bmatrix} \hat{\mu}_P^I - \mu_P \\ \hat{\mu}_P^{II} - \mu_P \\ \hat{\mu}_Q^I - \mu_Q \\ \hat{\mu}_Q^{II} - \mu_Q \end{bmatrix} \xrightarrow{d} N(\mathbf{0}, \begin{bmatrix} \frac{V_P}{n_I^{PQ}} & \mathbf{0} & \frac{V_{PQ}}{n_I^{PQ}} & \mathbf{0} \\ \mathbf{0} & \frac{V_P}{n_{II}^P} & \mathbf{0} & \mathbf{0} \\ \frac{V_{QP}}{n_I^{PQ}} & \mathbf{0} & \frac{V_Q}{n_I^{PQ}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{V_Q}{n_{II}^Q} \end{bmatrix}). \quad (29)$$

Hence, the Slutsky's theorem implies that:

$$\hat{\mu}_P - \hat{\mu}_Q \xrightarrow{d} N(\mathbf{0}, \mathbf{R} \begin{bmatrix} \frac{V_P}{n_I^{PQ}} & \mathbf{0} & \frac{V_{PQ}}{n_I^{PQ}} & \mathbf{0} \\ \mathbf{0} & \frac{V_P}{n_{II}^P} & \mathbf{0} & \mathbf{0} \\ \frac{V_{QP}}{n_I^{PQ}} & \mathbf{0} & \frac{V_Q}{n_I^{PQ}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{V_Q}{n_{II}^Q} \end{bmatrix} \mathbf{R}'). \quad (30)$$

Replacing \mathbf{R} , V_P , V_Q , V_{PQ} , and V_{QP} with their consistent estimators, Equation (30) implies that:

$$\begin{aligned} \hat{\mu}_P - \hat{\mu}_Q &\xrightarrow{d} N(\mathbf{0}, \hat{V}), \\ \text{where } \hat{V} &= (\frac{n_I^{PQ}}{n_I^{PQ}+n_{II}^P}, \frac{n_{II}^P}{n_I^{PQ}+n_{II}^P}, -\frac{n_I^{PQ}}{n_I^{PQ}+n_{II}^Q}, -\frac{n_{II}^Q}{n_I^{PQ}+n_{II}^Q}) \begin{bmatrix} \frac{\hat{V}_P}{n_I^{PQ}} & \mathbf{0} & \frac{\hat{V}_{PQ}}{n_I^{PQ}} & \mathbf{0} \\ \mathbf{0} & \frac{\hat{V}_P}{n_{II}^P} & \mathbf{0} & \mathbf{0} \\ \frac{\hat{V}_{QP}}{n_I^{PQ}} & \mathbf{0} & \frac{\hat{V}_Q}{n_I^{PQ}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{\hat{V}_Q}{n_{II}^Q} \end{bmatrix} \begin{bmatrix} \frac{n_I^{PQ}}{n_I^{PQ}+n_{II}^P} \\ \frac{n_{II}^P}{n_I^{PQ}+n_{II}^P} \\ -\frac{n_I^{PQ}}{n_I^{PQ}+n_{II}^Q} \\ -\frac{n_{II}^Q}{n_I^{PQ}+n_{II}^Q} \end{bmatrix} \\ &= \frac{\hat{V}_P}{n_I^{PQ}+n_{II}^P} + \frac{\hat{V}_Q}{n_I^{PQ}+n_{II}^Q} - \frac{n_I^{PQ}}{(n_I^{PQ}+n_{II}^P)(n_I^{PQ}+n_{II}^Q)}(\hat{V}_{PQ} + \hat{V}_{QP}). \end{aligned} \quad (31)$$

We conclude that $\hat{\mu}_P - \hat{\mu}_Q$ is a $K - 1$ dimensional vector that converges to a multivariate normal distribution with zero mean and \hat{V} variance. This implies that:

$$\hat{u}_p \equiv (\hat{\mu}_P - \hat{\mu}_Q)' \hat{V}^{-1} (\hat{\mu}_P - \hat{\mu}_Q) \xrightarrow{d} \chi^2(K - 1). \quad (32)$$

□

Theorem 5. *Under assumptions A1 to A3, if all individuals have Rational Expectations about Y_i (Definition 1), then*

$$E_i(y_i) = E_i(\mu_{it}^B) \text{ and } E_i(\text{var}(Y_{it}^B)) + \text{var}_i(\mu_{it}^B) = \text{var}_i(y_i).$$

Proof. The first part of Theorem 5 is the same as the second part of Theorem 1, which we have already proved above. Here we prove the second part of Theorem 5.

Recall from the proof of Theorem 1 that $\mu_{it}^B = E[Y_i | (\Theta_i = \theta_i, \mathbf{X}_i^{\theta_i, t^-} = \mathbf{x}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-} = \mathbf{z}^{t-})]$. As above, the population distribution of Y_{it}^B is produced by the realizations of Θ_i and $\mathbf{X}_i^{\theta_i, t^-}$. Hence, the term $\text{var}_i(\mu_{it}^B)$ is given by:

$$\text{var}_i(\mu_{it}^B) = \text{var}\{E[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t-})] | \mathbf{Z}^{t-} = \mathbf{z}^{t-}\}. \quad (33)$$

Theorem 1 states that $E_i(\text{var}(Y_{it}^B)) = E\{\text{var}[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t-})] | \mathbf{Z}^{t-} = \mathbf{z}^{t-}\}$. Hence, the sum of $E_i(\text{var}(Y_{it}^B))$ and $\text{var}_i(\mu_{it}^B)$ is given by:

$$\begin{aligned} E_i(\text{var}(Y_{it}^B)) + \text{var}_i(\mu_{it}^B) &= E\{\text{var}[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t-})] | \mathbf{Z}^{t-} = \mathbf{z}^{t-}\} + \\ &\quad \text{var}\{E[Y_i | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-}, \mathbf{Z}^{t-})] | \mathbf{Z}^{t-} = \mathbf{z}^{t-}\} \\ &= \text{var}(Y_i | \mathbf{Z}^{t-} = \mathbf{z}^{t-}) \\ &= \text{var}_i(y_i), \end{aligned} \quad (34)$$

where the second last line follows from the law of iterative variance. □

Theorem 6. (*Parametric Aggregate-Distribution Test*) *Under assumptions A1 to A3 and A5, if all individuals have Rational Expectations about Y_i (Definition 1), then*

$$\begin{aligned} \hat{u}_M &\equiv \frac{\hat{\mu}_{BM} - \hat{\mu}_{YM}}{\sqrt{\frac{\hat{\sigma}_{BM}^2}{n^B} + \frac{\hat{\sigma}_{YM}^2}{n^Y} - \frac{2n^{BY}\hat{\sigma}_M^2}{n^B n^Y}}} \xrightarrow{d} N(0, 1), \\ \hat{u}_V &\equiv \frac{\hat{\mu}_{BV} - \hat{\mu}_{YV}}{\sqrt{\frac{\hat{\sigma}_{BV}^2}{n^B} + \frac{\hat{\sigma}_{YV}^2}{n^Y} - \frac{2n^{BY}\hat{\sigma}_V^2}{n^B n^Y}}} \xrightarrow{d} N(0, 1). \end{aligned}$$

Proof. We note that, in terms of mathematical content, Theorem 6 is in essence a one dimensional special case of Theorem 4. Hence, the proof for Theorem 6 follows the same structure as the proof for Theorem 4. □

Theorem 7. Any two of the following three statements imply the third statement:

- (1) $\delta_{t,1}^B = \delta_{t,2}^B$, and $\sqrt{\delta_{t,3}^B} = 0$; (Mean-Variance Test)
- (2) $E_i(\text{var}(Y_{it}^B)) + \text{var}_i(\mu_{it}^B) = \text{var}_i(y_i)$, and $E_i(y_i) = E_i(\mu_{it}^B)$; (Aggregate-Distribution Test)
- (3) $y_i = \mu_{it}^B + \epsilon_i$, where ϵ_i has a mean of zero and is orthogonal to μ_{it}^B . (Regression Test)

Proof. Let v_{it}^B denote $\text{var}(Y_{it}^B)$. We start by rewriting the three statements as follows:

- (1) $E_i(v_{it}^B) = \text{var}_i(y_i - \mu_{it}^B)$, and $E_i(y_i) - E_i(\mu_{it}^B) = 0$;
- (2) $E_i(v_{it}^B) + \text{var}_i(\mu_{it}^B) = \text{var}_i(y_i)$, and $E_i(y_i) = E_i(\mu_{it}^B)$;
- (3) $\text{cov}(y_i - \mu_{it}^B, \mu_{it}^B) = 0$, and $E_i(y_i - \mu_{it}^B) = 0$.

It is easy to see that the second components of these statements are identical. Hence, it suffices to prove that any two of the following three equalities imply the third:

$$E_i(v_{it}^B) = \text{var}_i(y_i - \mu_{it}^B) \iff E_i(v_{it}^B) = \text{var}_i(y_i) + \text{var}_i(\mu_{it}^B) - 2\text{cov}(y_i, \mu_{it}^B), \quad (35)$$

$$E_i(v_{it}^B) + \text{var}_i(\mu_{it}^B) = \text{var}_i(y_i) \iff E_i(v_{it}^B) = \text{var}_i(y_i) - \text{var}_i(\mu_{it}^B), \quad (36)$$

$$\text{cov}(y_i - \mu_{it}^B, \mu_{it}^B) = 0 \iff \text{cov}(y_i, \mu_{it}^B) = \text{var}(\mu_{it}^B). \quad (37)$$

Taking the difference between Equation (35) and Equation (36), we find that $0 = 2\text{var}_i(\mu_{it}^B) - 2\text{cov}(y_i, \mu_{it}^B)$, which is the same as Equation (37).

Equation (37) implies that $\text{var}(\mu_{it}^B) - 2\text{cov}(y_i, \mu_{it}^B) = -\text{var}_i(\mu_{it}^B)$. Hence, Equation (35) and Equation (36) are identical if Equation (37) holds. □

Theorem 8. Under assumptions A1 and A6-A8, if all individuals have Rational Expectations about Y_i (Definition 1), then

$$\delta_{t,1}^B - \delta_{t,2}^B \geq 0.$$

Proof. Consider an individual i of type θ_i . Under the assumption that she has Rational Expectations about Y_i , her beliefs at t about y_i , Y_{it}^B , are given by:

$$\begin{aligned} Y_{it}^B &= Y_i | (\mathbf{X}_i^{\theta_i, t^-} = \mathbf{x}_i^{\theta_i, t^-}, \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}) \\ &= Y_i^I | \mathbf{X}_i^{\theta_i, t^-} = \mathbf{x}_i^{\theta_i, t^-} + Y^A | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}, \end{aligned} \quad (38)$$

where the second line follows from the assumptions that Θ_i , \mathbf{X}_i , and \mathbf{Z} are independent (A7).

Under these independence assumptions, $\Theta_i = \theta_i$ does not provide additional information

about Y_i^I , conditional on $\mathbf{X}_i^{\theta_i, t^-}$. Hence, the following equation holds:

$$Y_{it}^B = Y_i^I | (\Theta_i = \theta_i, \mathbf{X}_i^{\theta_i, t^-} = \mathbf{x}_i^{\theta_i, t^-}) + Y^A | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}. \quad (39)$$

Taking the mean and variance of both sides, the independence assumptions imply that:

$$\mu_{it}^B = E(Y_{it}^B) = E[Y_i^I | (\Theta_i = \theta_i, \mathbf{X}_i^{\theta_i, t^-} = \mathbf{x}_i^{\theta_i, t^-})] + E(Y^A | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}), \quad (40)$$

$$\text{var}(Y_{it}^B) = \text{var}[Y_i^I | (\Theta_i = \theta_i, \mathbf{X}_i^{\theta_i, t^-} = \mathbf{x}_i^{\theta_i, t^-})] + \text{var}(Y^A | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}). \quad (41)$$

The population distribution of Y_{it}^B is produced by the realizations of Θ_i and $\mathbf{X}_i^{\theta_i, t^-}$. Hence, the first component, $\delta_{t,1}^B = E_i[\text{var}(Y_{it}^B)]$, is given by:

$$\delta_{t,1}^B = E\{\text{var}[Y_i^I | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})]\} + \text{var}(Y^A | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-}). \quad (42)$$

Similarly, the population distribution of y_i is produced by the realizations of \mathbf{X}_i . Hence, the second component, $\delta_{t,2}^B = \text{var}_i(y_i - \mu_{it}^B)$, is given by:

$$\delta_{t,2}^B = \text{var}_{(\Theta_i, \mathbf{X}_i)}\{Y_i^I - E[Y_i^I | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})] + Y^A | \mathbf{Z} = \mathbf{z} - E(Y^A | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-})\}. \quad (43)$$

Note that $Y^A | \mathbf{Z} = \mathbf{z} - E(Y^A | \mathbf{Z}^{t^-} = \mathbf{z}^{t^-})$ is a constant across individuals, hence it does not contribute to the cross-sectional variance of $y_i - \mu_{it}^B$. It implies that:

$$\begin{aligned} \delta_{t,2}^B &= \text{var}\{Y_i^I - E[Y_i^I | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})]\} \\ &= \text{var}\{E\{Y_i^I - E[Y_i^I | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})] | \Theta_i\}\} + E\{\text{var}\{Y_i^I - E[Y_i^I | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})] | \Theta_i\}\} \\ &= \text{var}(0) + E\{\text{var}\{Y_i^I - E[Y_i^I | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})] | \Theta_i\}\} \\ &= E\{\text{var}\{E\{Y_i^I - E[Y_i^I | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})] | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})\} | \Theta_i\}\} + \\ &\quad E\{E\{\text{var}\{Y_i^I - E[Y_i^I | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})] | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})\} | \Theta_i\}\} \\ &= E[\text{var}(0 | \Theta_i)] + E\{E\{\text{var}[Y_i^I | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})] | \Theta_i\}\} \\ &= E\{\text{var}[Y_i^I | (\Theta_i, \mathbf{X}_i^{\theta_i, t^-})]\}. \end{aligned} \quad (44)$$

where the second and fourth lines follow from the law of iterative variance.

Taking the difference between $\delta_{t,1}^B$ and $\delta_{t,2}^B$, we conclude:

$$\delta_{t,1}^B - \delta_{t,2}^B = \text{var}(Y^A | \mathbf{Z}^{t-} = \mathbf{z}^{t-}) \geq 0. \quad (45)$$

□

Theorem 9. *Under assumptions A1 and A6-A8, if all individuals have Rational Expectations about Y_i (Definition 1), then Y_t^B is a mean-preserving spread of $Y_z + E(Y_t^B) - E(Y_z)$.*

Proof. Under Rational Expectations, the subjective aggregate distribution Y_t^B is given by:

$$\begin{aligned} Y_t^B &\equiv E_i(Y_{it}^B) \\ &= E_i[Y_i^I | (\Theta_i = \theta_i, \mathbf{X}_i^{\theta_i, t-} = \mathbf{x}_i^{\theta_i, t-}) + Y^A | \mathbf{Z}^{t-} = \mathbf{z}^{t-}] \\ &= E[Y_i^I | (\Theta_i, \mathbf{X}_i^{\theta_i, t-})] + Y^A | \mathbf{Z}^{t-} = \mathbf{z}^{t-} \\ &= Y_i^I + Y^A | \mathbf{Z}^{t-} = \mathbf{z}^{t-}, \end{aligned} \quad (46)$$

where the third line follows from the fact that cross-sectional variation in Y_{it}^B is produced by the realizations of Θ_i and $\mathbf{X}_i^{\theta_i, t-}$, and the fourth line follows from the law of total probability.

The distribution of outcomes Y_z is given by:

$$\begin{aligned} Y_z &\equiv Y_i | \mathbf{Z} = \mathbf{z} \\ &= Y_i^I + Y^A | \mathbf{Z} = \mathbf{z}, \end{aligned} \quad (47)$$

where the second equality follows from the independence of \mathbf{X}_i and \mathbf{Z} (A7).

Then, $Y_z + E(Y_t^B) - E(Y_z)$ is given by:

$$\begin{aligned} Y_z + E(Y_t^B) - E(Y_z) &= Y_i^I + Y^A | \mathbf{Z} = \mathbf{z} + E(Y_i^I + Y^A | \mathbf{Z}^{t-} = \mathbf{z}^{t-}) - E(Y_i^I + Y^A | \mathbf{Z} = \mathbf{z}) \\ &= Y_i^I + E(Y^A | \mathbf{Z}^{t-} = \mathbf{z}^{t-}). \end{aligned} \quad (48)$$

Hence, it is easy to verify that Y_t^B is equal to the sum of $Y_z + E(Y_t^B) - E(Y_z)$ and $Y^A | \mathbf{Z}^{t-} = \mathbf{z}^{t-} - E(Y^A | \mathbf{Z}^{t-} = \mathbf{z}^{t-})$. Since $Y^A | \mathbf{Z}^{t-} = \mathbf{z}^{t-} - E(Y^A | \mathbf{Z}^{t-} = \mathbf{z}^{t-})$ has a mean of zero, we conclude that Y_t^B is a mean-preserving spread of $Y_z + E(Y_t^B) - E(Y_z)$. □

Theorem 10. *For any combination of Y_z , which is continuous and has a finite expectation, and $E_i(P_{it}^{B,k})$, $k = 1, 2, \dots, K$, if 1) $\bigcup_{k=1}^K S_k = \mathbb{R}$ and 2) $E_i(P_{it}^{B,k}) > 0 \ \forall k$, then it is always possible to find a random variable Ψ and a constant c such that:*

- (1) Ψ is a mean-preserving spread of $Y_z + c$,
(2) $E_i(P_{it}^{B,k}) = \text{Prob}(\Psi \in S_k)$, for all k .

Proof. We prove this theorem by construction.

By assumption, there are K outcome bins, S_k . Among them, S_1 and S_K are unbounded and all other outcome bins are bounded. Let $f_\Psi^c(\psi)$ be an arbitrary positive-valued function such that $\int_{S_k} f_\Psi^c(\psi) d\psi = E_i(P_{it}^{B,k})$ for $k = 2, 3, \dots, K-1$. A simple example of such $f_\Psi^c(\psi)$ is a step function that takes the value $\frac{E_i(P_{it}^{B,k})}{\sup_{S_k} S_k - \inf_{S_k} S_k}$ for $\psi \in S_k$. Let μ_Ψ^c denote $\int_{\bigcup_{k=2}^{K-1} S_k} f_\Psi^c(\psi) \psi d\psi$ and let $F_\Psi^c(\psi)$ denote $E_i(P_{it}^{B,1}) + \int_{\inf S_2}^\psi f_\Psi^c(\psi') d\psi'$.

We let $c = \mu_\Psi^c - E(Y_z)$ and let Y_z^c denote $Y_z + c$. The PDF and CDF of Y_z^c are denoted as $f_{Y_z^c}(y)$ and $F_{Y_z^c}(y)$, respectively. Because Y_z^c is a continuous random variable, there exist unique y_l^* and y_r^* such that $F_{Y_z^c}(y_l^*) = E_i(P_{it}^{B,1})$ and $F_{Y_z^c}(y_r^*) = 1 - E_i(P_{it}^{B,K})$. We define $f_\Psi^l(\psi; W_l) = f_{Y_z^c}(\psi + W_l)$ for $\psi \leq y_l^* - W_l$ and $f_\Psi^r(\psi; W_r) = f_{Y_z^c}(\psi - W_r)$ for $\psi \geq y_r^* + W_r$.

Let μ_Ψ^l denote $\int_{-\infty}^{y_l^* - W_l} f_\Psi^l(\psi; W_l) \psi d\psi$ and let μ_Ψ^r denote $\int_{y_r^* + W_r}^\infty f_\Psi^r(\psi; W_r) \psi d\psi$. It implies that $\mu_\Psi^l = \int_{-\infty}^{y_l^*} f_{Y_z^c}(\psi) \psi d\psi - W_l E_i(P_{it}^{B,1})$ and $\mu_\Psi^r = \int_{y_r^*}^\infty f_{Y_z^c}(\psi) \psi d\psi + W_r E_i(P_{it}^{B,K})$. We note that both $\int_{-\infty}^{y_l^*} f_{Y_z^c}(\psi) \psi d\psi$ and $\int_{y_r^*}^\infty f_{Y_z^c}(\psi) \psi d\psi$ are finite because Y_z^c has finite first moment. Hence, for any W_l , there exists a unique W_r such that $\mu_\Psi^l + \mu_\Psi^r = 0$. We denote this unique mapping as $W_r^*(W_l) \equiv [W_l E_i(P_{it}^{B,1}) - \int_{-\infty}^{y_l^*} f_{Y_z^c}(\psi) \psi d\psi - \int_{y_r^*}^\infty f_{Y_z^c}(\psi) \psi d\psi] / [E_i(P_{it}^{B,K})]$.

We let W_l be an arbitrary real number such that 1) $W_l > \frac{\int_{y_l^*}^{y_r^*} F_{Y_z^c}(y) dy}{E_i(P_{it}^{B,1})}$, 2) $y_l^* - W_l < \inf S_2$, 3) $W_r^*(W_l) > 0$, and 4) $y_r^* + W_r^*(W_l) > \sup S_{K-1}$. It is easy to verify that all these constraints are finite lower bounds for W_l . Hence, since W_l is not bounded, there always exists such a W_l that is large enough to satisfy all the constraints.

Now we construct a random variable Ψ as follows. Let $F_\Psi(\psi)$ denote the CDF of Ψ . It is then given by:

$$F_\Psi(\psi) = \begin{cases} F_{Y_z^c}(\psi + W_l), & \text{if } \psi < y_l^* - W_l \\ E_i(P_{it}^{B,1}), & \text{if } y_l^* - W_l \leq \psi < \inf S_2 \\ F_\Psi^c(\psi), & \text{if } \inf S_2 \leq \psi < \sup S_{K-1} \\ 1 - E_i(P_{it}^{B,K}), & \text{if } \sup S_{K-1} \leq \psi < y_r^* + W_r^*(W_l) \\ F_{Y_z^c}(\psi - W_r^*(W_l)), & \text{if } y_r^* + W_r^*(W_l) \leq \psi \end{cases}. \quad (49)$$

We claim that Ψ satisfies the conditions that 1) Ψ is a mean-preserving spread of Y_z^c , and 2) $E_i(P_{it}^{B,k}) = \text{Prob}(\Psi \in S_k)$, for all k . It is easy to verify that 2) holds by construction, so

we will focus on proving 1) below.

To prove 1), it suffices to show that a) $E(\Psi) = E(Y_z^C)$ and b) $\int_{-\infty}^x F_{\Psi}(x') - F_{Y_z^C}(x') dx' \geq 0$ for all x . Note that $E(\Psi) = \mu_{\Psi}^l + 0 + \mu_{\Psi}^c + 0 + \mu_{\Psi}^r = \mu_{\Psi}^c$ and $E(Y_z^C) = E(Y_z) + c = E(Y_z) + \mu_{\Psi}^c - E(Y_z) = \mu_{\Psi}^c$. Hence, a) is satisfied.

For any $x < y_l^*$, we know that $F_{\Psi}(x) - F_{Y_z^C}(x) = \min(F_{Y_z^C}(x + W_l), F_{Y_z^C}(y_l^*)) - F_{Y_z^C}(x) \geq 0$ because $F_{Y_z^C}(x)$ is monotonically increasing. Hence, $\int_{-\infty}^x F_{\Psi}(x') - F_{Y_z^C}(x') dx' \geq 0$ for any $x < y_l^*$.

Consider $y_l^* \leq x < y_r^*$. We have:

$$\begin{aligned}
\int_{-\infty}^x F_{\Psi}(x') dx' &= \int_{-\infty}^{y_l^* - W_l} F_{\Psi}(x') dx' + \int_{y_l^* - W_l}^{y_l^*} F_{\Psi}(x') dx' + \int_{y_l^*}^x F_{\Psi}(x') dx' \\
&= \int_{-\infty}^{y_l^*} F_{Y_z^C}(x') dx' + W_l E_i(P_{it}^{B,1}) + \int_{y_l^*}^x F_{\Psi}(x') dx' \\
&> \int_{-\infty}^{y_l^*} F_{Y_z^C}(x') dx' + \int_{y_l^*}^{y_r^*} F_{Y_z^C}(y) dy + 0 \\
&\geq \int_{-\infty}^{y_l^*} F_{Y_z^C}(x') dx' + \int_{y_l^*}^x F_{Y_z^C}(y) dy \\
&= \int_{-\infty}^x F_{Y_z^C}(x') dx'.
\end{aligned} \tag{50}$$

Lastly we consider $x \geq y_r^*$. We have shown above that $\int_{-\infty}^{y_r^*} F_{\Psi}(x') - F_{Y_z^C}(x') dx' > 0$. Since $E(\Psi) = E(Y_z^C)$, we know that $\int_{-\infty}^{\infty} F_{\Psi}(x') - F_{Y_z^C}(x') dx' = 0$. Note that for any $x \geq y_r^*$, we have $F_{\Psi}(x) - F_{Y_z^C}(x) = \max(F_{Y_z^C}(x - W_r^*(W_l)), F_{Y_z^C}(y_r^*)) - F_{Y_z^C}(x) \leq 0$. This implies that $\int_{-\infty}^x F_{\Psi}(x') - F_{Y_z^C}(x') dx'$ is monotonically decreasing for $x \geq y_r^*$. Therefore, we conclude that $\int_{-\infty}^x F_{\Psi}(x') - F_{Y_z^C}(x') dx' > 0$ for all $x \geq y_r^*$.

Combining the arguments above, we have shown that $\int_{-\infty}^x F_{\Psi}(x') - F_{Y_z^C}(x') dx' > 0$ for all x . This concludes the proof for Theorem 10. \square

E Existing Mean-based Tests

We apply the existing mean-based tests (DGM, Regression and Correlation) to beliefs about income at age 28 (or 29). DGM and Regression Tests only utilize data on beliefs and realizations. In contrast, since the Correlation Test concerns a strong version of Rational Expectations Hypothesis, this test requires specifying variables in the agent's information set at the time of belief elicitation. Given the strong serial correlation of income documented in

the literature, one piece of information that respondents should use to predict future income is current income. For the post-college period, we first compute the “innovation”, which is the difference between realized annual family income at age 28 (or 29) and the mean of beliefs about family income at 28 (or 29) elicited 1 or 5 years earlier. We then compute the cross-sectional correlation between this innovation and realized annual family income in the year when beliefs are elicited, and test whether this correlation equals zero. Many students have little income, so for the in-school period we considered the correlation between the innovation in own income (the difference between beliefs and realized income) and the cumulative college GPA at the time of belief elicitation.

Detailed results for these tests are shown in Table 9. For the in-school period, the three tests produce rejections of the null at a 5% level for almost every year.¹⁷ In contrast, in the post-college period these tests reject much less frequently. In particular, the DGM test does not reject at conventional significance levels; the Regression test rejects for 5-years ahead beliefs but not 1-year ahead; and the correlation test rejects for 1-year ahead beliefs but not 5-year ahead.

Table 9: Rational Expectations Tests about Income at Age 28 (29)

Year of Belief-elicitation	10-year ahead	9-year ahead	8-year ahead	7-year ahead	6-year ahead	5-year ahead	1-year ahead
DGM Test (pooled data): p-value	0.002	$< 10^{-4}$	0.0100	0.0180	0.0581	0.9840	0.3928
DGM Test (balanced panel): p-value	0.0120	0.0140	0.0301	0.0361	0.0922	0.3066	0.7916
Regression Test: Intercept	\$24,140	\$19,366	\$22,032	\$20,938	\$13,098	\$11,622	\$480
Regression Test: Slope	0.0884	0.1937	0.1597	0.2037	0.4379	0.6822	0.9956
Regression Test: p-value	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	0.0006	0.9276
Correlation Test: Correlation	N.A.	0.2109	0.2283	0.2165	0.1236	-0.0473	0.1644
Correlation Test: p-value	N.A.	0.0006	0.0007	0.0016	0.0769	0.5574	0.0050

Note: All tests are applied to balanced panels. The first five columns report the results for existing mean-based tests described in Section 4.4 for students’ beliefs in college about own annual income at age 28. The last two columns report the results for existing mean-based tests for students’ beliefs after college about annual family income at age 28 (or 29).

¹⁷The only exceptions are the DGM Test and the Correlation Test for end-of-college income beliefs, where the null of Rational Expectations are rejected at a 10% level.

F Nonparametric Test of Rational Expectations with Aggregate Shocks: an Example

We consider the case where $\bigcup_{k=1}^K S_k = \mathbb{R}$ is finite. Specifically, we consider the following example:

Example 1. *There are three outcome bins, which are $S_1 = (-5, -0.5]$, $S_2 = (-0.5, 0.5]$, and $S_3 = (0.5, 5]$. The distribution of Y_i is $N(1, 1)$ truncated to $(-2.5, 4.5)$. The aggregate perceived probabilities are given by $E_i(P_{it}^{B,1}) = E_i(P_{it}^{B,3}) = 0.005$, $E_i(P_{it}^{B,2}) = 0.99$.*

In Example 1, because adding a constant c to random variable Y_i does not change its variance, the variance of $Y_i + c$ is 0.9939 for all c . The aggregate subjective distribution is concentrated on S_2 . Below we derive a bound for the variance of any distribution that is consistent with the observed aggregate perceived probabilities in this example.

Consider an arbitrary random variable Ψ such that $Prob(\Psi \in S_1) = Prob(\Psi \in S_3) = 0.005$ and $Prob(\Psi \in S_2) = 0.99$. Let μ^Ψ denote $E(\Psi)$ and μ_k^Ψ denote $E(\Psi | \Psi \in S_k)$ for $k = 1, 2, 3$. The variance of Ψ is given by:

$$\begin{aligned}
 var(\Psi) &= E(\psi - \mu^\Psi)^2 \\
 &= E(\psi - \mu_2^\Psi)^2 + E(\psi - \mu^\Psi)^2 - E(\psi - \mu_2^\Psi)^2 \\
 &= E(\psi - \mu_2^\Psi)^2 + E(2\psi - \mu^\Psi - \mu_2^\Psi)(\mu_2^\Psi - \mu^\Psi) \\
 &= E(\psi - \mu_2^\Psi)^2 - (\mu_2^\Psi - \mu^\Psi)^2 \\
 &\leq 0.99E[(\psi - \mu_2^\Psi)^2 | \Psi \in S_2] + 0.005E[(\psi - \mu_2^\Psi)^2 | \Psi \in S_1] + 0.005E[(\psi - \mu_2^\Psi)^2 | \Psi \in S_3] \\
 &\leq 0.99E[(\psi - \mu_2^\Psi)^2 | \Psi \in S_2] + 0.005(-5 - 0.5)^2 + 0.005[5 - (-0.5)]^2 \\
 &\leq 0.99 \frac{[0.5 - (-0.5)]^2}{4} + 0.1512 \\
 &= 0.3987,
 \end{aligned} \tag{51}$$

where the second last line follows from Popoviciu's inequality.

Hence, we conclude that any random variable that is consistent with the observed aggregate perceived probabilities must have a variance that is smaller than the variance of $Y_i + c$. It implies that it is impossible to find a pair of Ψ and c such that both (1) and (2) of Theorem 9 are satisfied. Thus, we reject Rational Expectations for Example 1.