How to Collect Your Own Data: A Primer on Survey Design

Christopher Rauh
University of Cambridge

2024

Why use surveys?

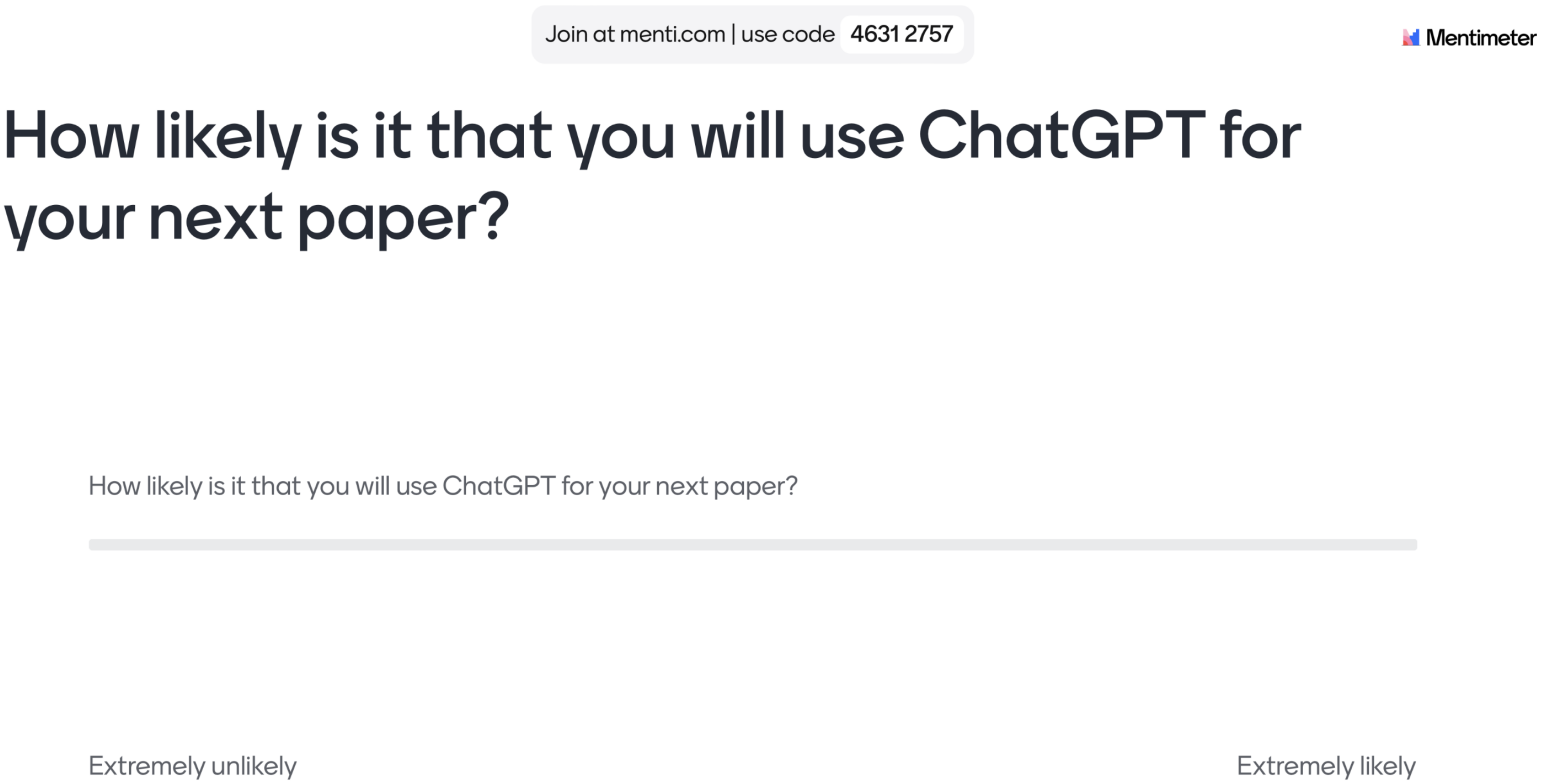
- Many things are invisible in administrative data or conventional surveys: perceptions, attitudes, knowledge, and views.
- You can create your own controlled and identifying variation.
- Surveys need to be well-designed, carefully calibrated, and deployed on appropriate samples.

Disclaimer: Much (but not all) of what you will see today stems from "How to run surveys: A guide to creating your own identifying variation and revealing the invisible" (Stantcheva 2022). For the sake of exposition I am not flagging her excellent material every time it has been incorporated.

Let us start with an example

- Go to www.menti.com
- Enter code 4631 2757

Survey example



How likely is it that you will use ChatGPT for your next paper?

Is this a "good" question?

- Am I asking about a probability?
- What do I mean by "using for your next paper"?
- You might be afraid that I mean "cheating"....
  - In this case: Will you tell the truth?

Why study perceptions/beliefs?

- "Traditional" models assume:
  - Everybody knows the skills production function
  - Everybody knows the returns to skills
- "Fundamentalists" would go as far as saying:
  - There is no heterogeneity in preferences
- So how can we reconcile some of the differences in investments we see?
  - Is it all about constraints?

Why not rely on observational data alone?

- There exists a literature that uses observed choices to infer beliefs from heterogeneity in choices.
  - The idea is that all deviations from the "optimal" choice must be driven by differences in beliefs.
- Is this a good idea?
  - Investment choices are consistent with many different alternative specifications of preferences (Manski 2004).

Why might beliefs matter for policy?

Examples relating to education investments:

- If parents think school quality and own investments are complements, then by improving schools, parents invest more as well
- If parents think university is what matters for child development then policy might have to compensate for underinvestments in early childhood
- If minority students believe consumption value of university is low, then reducing tuition fees might not increase their enrollment
- .....

What we will see in this workshop – Surveys

- How to collect survey data.
- How to ask questions.
- Practicalities and considerations when dealing with survey companies.

What we will might see in this workshop – Research examples

- How to motivate the usage of beliefs data.
- Different contexts in which can measure beliefs.
- Some limitations of relying on surveys:
  - Beliefs: Where do they come from? Belief formation? Motivated beliefs? Are the beliefs correct?
  - Experiments: Is the effect "real"?

Surveys as a research tool

Online surveys

1. People can take surveys at their convenience
2. Mobile technologies are convenient, so may encourage some people who would not otherwise want to fill out questionnaires or answer questions on the phone to take surveys
3. Can reach people otherwise difficult to reach (e.g., people in remote areas, people who move a lot)
4. Offer a variety of rewards, so can appeal to a broader group of people
5. No need to train (and trust) enumerators
6. They are much cheaper!!!!

Main steps

1. Sample.
2. Length.
3. Focus on the main aim:
   - Measurement of beliefs.
   - Experiment (information treatment).
4. Optimization.

Sample options

- Commercial survey companies which use quota sampled panels (e.g., Qualtrics, Pureprofile).
- Commercial survey marketplaces (e.g., Lucid).
- Piggy-backing on large panels by including a module (e.g., SOEP, Understanding Society)
- Asking people in the street or by mass email

Trade-offs:

- Feasibility (e.g., legal consent from those underaged).
- Costs.
- Does sample relate to research question?
- Generalizable? You might be interested in certain demographic group (young men in the East) but referees tend to prefer "representative" populations

Online Survey Providers (Haaland et al 2023)

Key Providers

- Dynata: Formerly known as Research Now and Survey Sampling International.
- Lucid: Approximates marginal distributions of basic demographics.
- YouGov: Matches higher-dimensional cells of the population (e.g., age $\times$ gender).
  - Advantages: Flexibility in study design, including obfuscated follow-up studies.
  - Disadvantages: Self-selection bias may differ from the broader population.

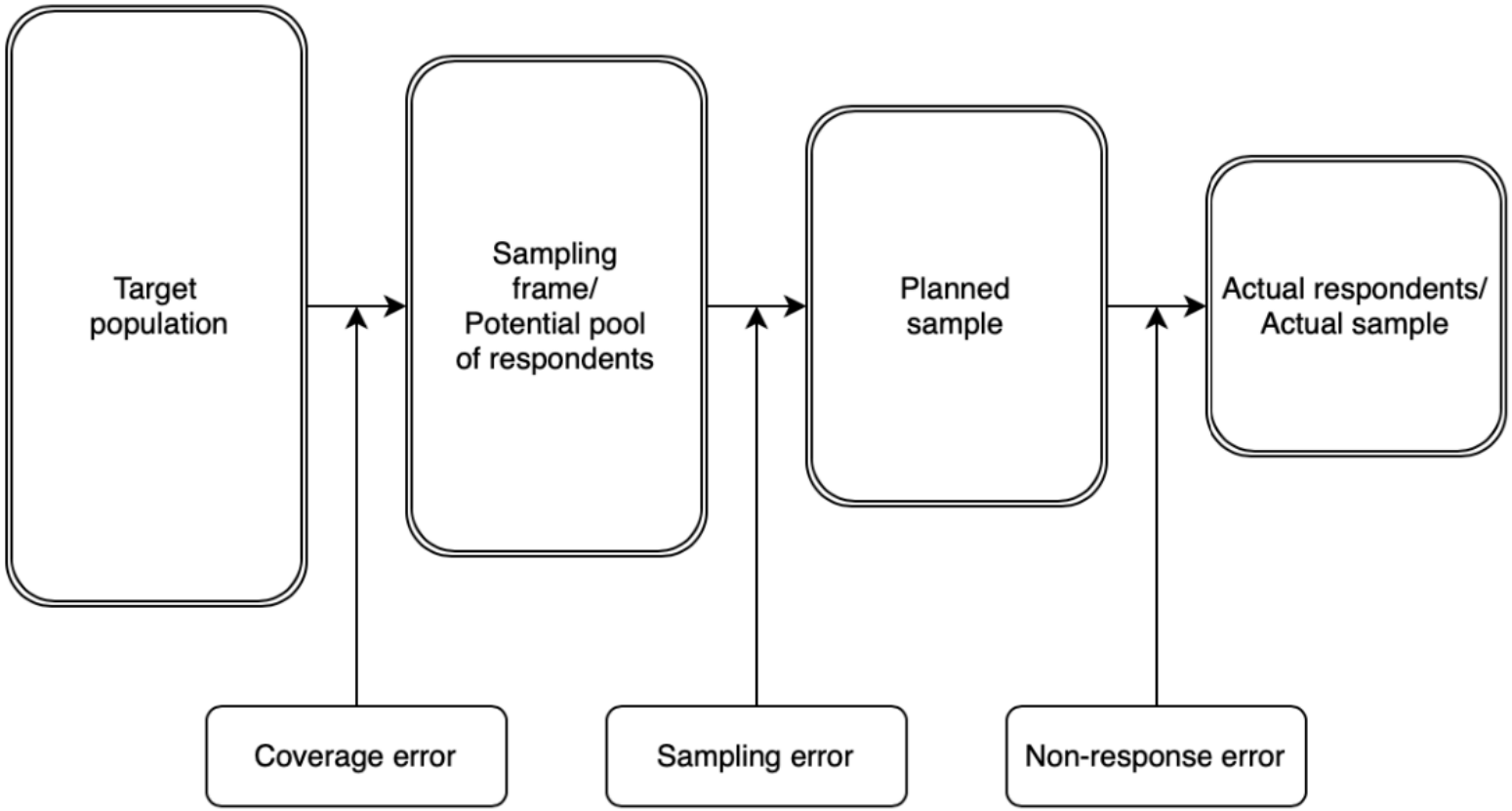
Research Findings

- Grewenig et al. (2018): Minimal differences between online and offline populations in political views, after controlling for survey method and respondent characteristics.
- Coppock and McClellan (2019): Lucid samples are comparable to ANES and GSS on Big Five personality, political knowledge, and framing effects.
- Haaland and Roth (2023): Similar experimental results with both representative online panels and probability-based samples.

Populations and sampling

- Target population: population of interest
- Sampling frame: all the people in the population you can potentially sample
- Coverage error: difference between sampling frame and the target population. For example, in online surveys, cannot survey people who are not online
- Planned sample: all people you want to complete your survey
- Sampling error: difference between planned sample and the sampling frame
- Actual sample: the people who end up taking your survey
- Non-response error: difference between the target sample and the actual sample. E.g., due to not seeing invite
- unit non-response bias: difference between respondents who start survey and those in planned sample
- item non-response: when respondents start survey but some answers are missing. Includes attrition – respondents dropping out before finishing.

Selection into survey



Source: Stantcheva (2022)

Targeting distributions

- Marginals:
  - 50% men, 50% women. 20% uni, 80% no uni.
  - It could happen that all university graduates are women.
  - It could happen that the final cell(s) you want to fill don't exist. (e.g. Women without university degree in Manhatten)
- Joint distribution:
  - 8% men with uni, 12% women with uni, 42% men without uni, 38% women without uni.
  - You might have problems locating the white women with three children and no college education in New York.
- My take: Start with joint, and then when progress stalls, switch to marginals. More likely to have good balance.

National representativeness of online surveys

- In the US and other high-income countries,
  - Online samples can offer good representation of a broad income spectrum ($25,000-100,000$), but very poor or rich are missing
  - Respondents skew more educated, more white, and somewhat more Democratic
  - Respondents from larger urban areas and urban clusters are overrepresented
- Important to critically assess sample in light of survey methods and topic before generalizing
- In particular, non-probability sampling (e.g., the quota sampling by survey company) has risks in terms of representativeness

If recruiting privately (e.g. school, firm, ...)

- Convince institution that you will not create extra work for them and that there are no potential reputational damages.
  - People are risk averse and have little to gain by participating.
    - ★ Try to convince them that they have something to gain (besides contributing to science). For instance, suggest you will send them personalized report.
- These points are important for any data you are ever going to want from anybody.
- Once survey link is shared with participants, incentivize through lottery rather than piece rates if you have a small budget.

Length

- Keep it short!
  - People get tired.
  - It's cheaper.
- Forcing responses can be useful, but can also be tiring. Use with care.
- It requires a lot of discipline to keep it short as it is tempting to ask everything.
  - Question every question. Will I use the information? How will I use it?

Design

Think (worry) about attention span of respondents:

- Avoid cluttering pages.
  - People get tired.
  - People get distracted and confused.
- Check design on your phone and computer with difference browsers
  - Make sure all questions are clearly visible.
  - Make sure "continue" button is clearly visible
- It requires a lot of discipline to keep it short as one is tempted to want to ask everything.

Open text

- Open-ended questions: have an open answer field of varying lengths (e.g., "What comes to mind when you think about income taxes?" [empty text field])
  - Natural Language Processing (NLP) is finally becoming common feature in Economics. There are many ways of quantitatively evaluating written text.
    - ★ Topic models (LDA). Embeddings (Bert). See Ash and Hansen (2022).
  - Think carefully when designing this question.
    - ★ We once asked people why they want to go to university and many responded why they might not. Think carefully!
- Hybrid questions: are close-ended questions with an open-ended answer choice (e.g., "Other (please specify):" [empty text field])
  - Less useful for analysis but can give you an idea why people are choosing "other".

Quantitative versus Likert scale

- Likert scales: "Strongly disagree" to "strongly agree" or "very unlikely" to "very likely" or "never" to "always"
  - Advantage: Intuitive and accessible for respondents
  - Disadvantage: How to compare across people and assess quantitatively?
- Probabilistically: Elicit percentage probability
  - Advantage: Comparable across people and can be assessed quantitatively
  - Disadvantage: Assumes numerical literacy, which can be especially tricky for conditional versus unconditional probabilities
- When using probabilistic questions, start with warm up question: "How likely do you think it is that it is going to rain today?"
- Other issue: Questions related to utility (e.g., "How much will you enjoy university?")
  - Can be sidestepped by asking: "How likely is it that you are going to enjoy university"?

Multiple Measurements of Beliefs

- Cognitive strain may induce measurement error in belief assessments.
- **Instrumental Variables (IV) Approach:**
  - Proposed by Gillen, Snowberg, and Yariv (2019) to mitigate classical measurement error.

Multiple Measurements in Belief Research

- (i) A qualitative survey question.
- (ii) A quantitative point estimate.
- (iii) A probabilistic question with mutually exclusive states.

Application and Considerations

- Example: Giglio et al. (2021) applied this IV approach to survey expectations about stock returns.
- Considerations:
  - Multiple measurements may be cognitively taxing, increasing survey fatigue or attrition rates.
  - Not applicable for nonclassical measurement error.

The main dependent variable

- Think about your main specification/model
- Probabilities give meaningful variation
  - But are a nightmare as a dependent variable, especially if a lot of 0% and 100% responses
    - ★ Otherwise could use fractional logit (which is not great to begin with)
    - ★ OLS usually does a good job but you will likely (rightfully?) face a referee who will complain
  - Tends to lead to some bunching
  - Some methods can deal with these sorts of distributions but they are often not familiar to referees.
- Binary variables can be convenient for all sorts of estimation techniques (e.g. maximum likelihood)
  - You lose lots of meaningful variation and cannot plot nice distributions
- Think carefully! Ideally pre-run your analysis with fake data in order to decide.

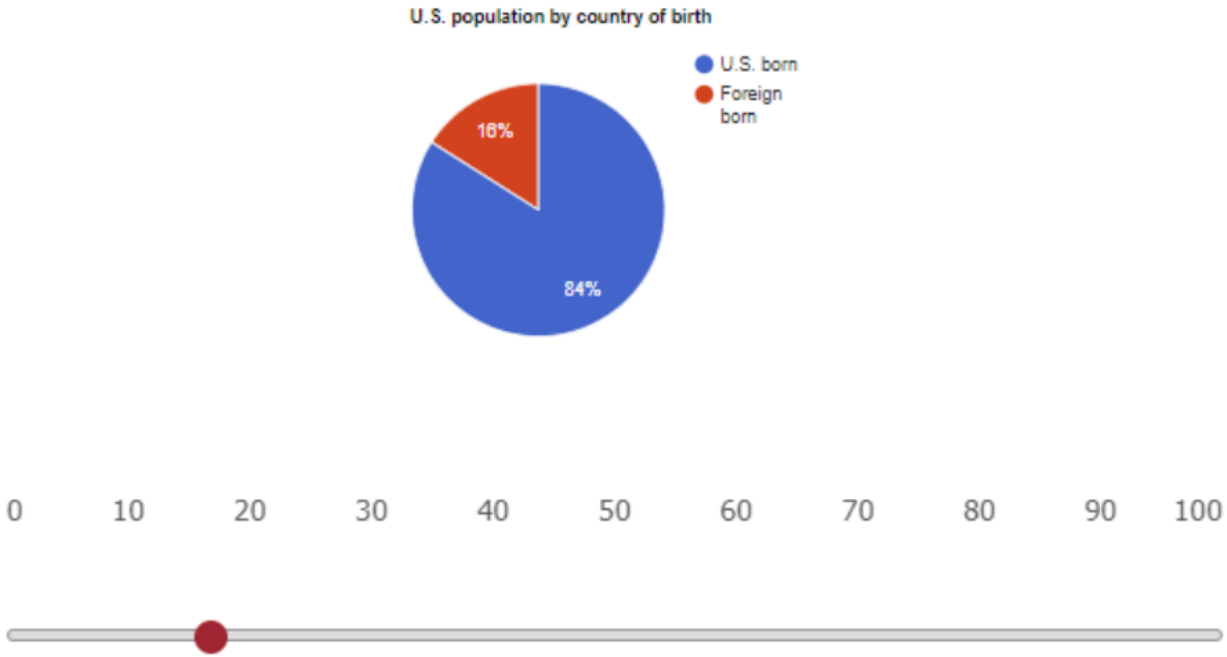
Scales

- Use natural metrics.
  - For example, use "last week" or "last month" rather than "regularly" or "often" when asking about how many times a respondent did something
- Label all options in an answer scale, not just extremes.
- Order of the response options.
  - Respondents may tend to pick last or first answer, so randomize order options that do not have natural ordering
- Remove numerical labels unless they have true meaning.
  - For example, do not add numeric labels to answer options such as "1 = strongly agree," "2 = agree"... "5 = strongly disagree."
  - This is distracting and potentially misleading
- Provide scales that approximate the actual distribution in the population or use open-ended questions to avoid biasing responses.
- Logically order answer options.

Sliders

Example slider question: share of immigrants from Alesina et al. (2022)

The pie chart below represents all the people currently living in the U.S. Out of all these people currently living in the U.S., how many do you think are legal immigrants? Move the slider to indicate how many out of every 100 people you think are legal immigrants.



Unclear: What should default value be on slider?

Complex questions

- Many studies will require new, creative, and sometimes complex questions.
  - For example, respondents may be asked to select their position on a ladder representing the income distribution using a slider.
- For these questions:
  - should pilot them them several times and check for understanding. Trust others, but also yourselves.
  - automate your survey code so that, e.g., percentages add up to 100.
    - ★ But be careful, residual category might contain lots of measurement error.
  - good visual representation is key
- Point beliefs vs probabilistic beliefs.
  - Point beliefs are easier for respondents to understand, but do not allow respondents to express uncertainty.
  - Probabilistic beliefs may be harder to understand and yield noisier results, but allow respondents to express uncertainty.

Example of complex question

Example of a complex question with visual representation from Alesina, Stantcheva, and Teso (2018)

We would now like to ask you what you think about the life opportunities of children from very poor families.

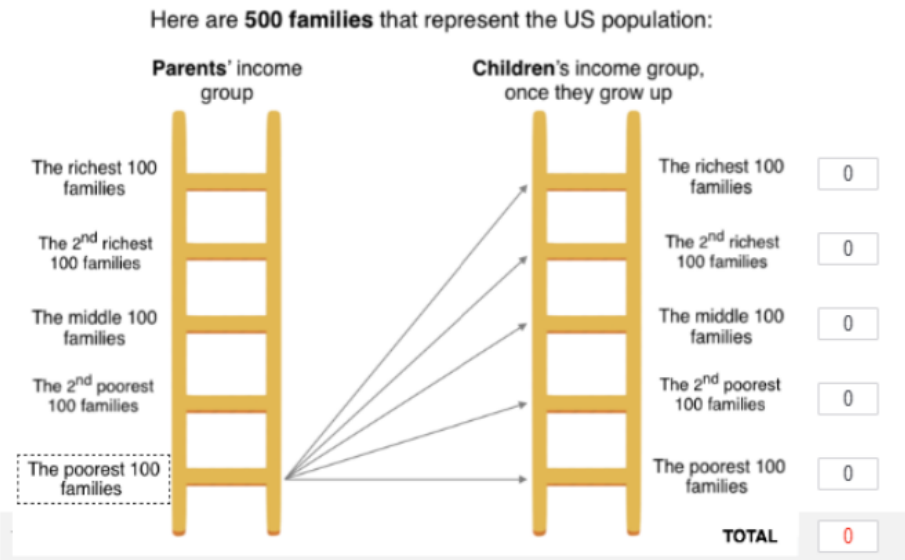
For the following questions, we focus on 500 families that represent the U.S. population. We divide them into five groups on the basis of their income, with each group containing 100 families. These groups are: the poorest 100 families, the second poorest 100 families, the middle 100 families, the second richest 100 families, and the richest 100 families.

In the following questions, we will ask you to evaluate the chances that children born in one of the poorest 100 families, once they grow up, will belong to any of these income groups.

Please fill out the entries to the right of the figure below to tell us, in your opinion, how many out of 100 children coming from the **poorest** 100 families will grow up to be in each income group.

From our experience, this question will take you at the very least 1 minute to answer.

Please note that your entries need to add up to 100 or you will not be able to move on to the next page.



Sometimes better to ask "out of 100 people how many" rather than "what is the probability that a person" will go to university

Anchoring

- In "How people update beliefs about climate change: Good news and bad news", Sunstein et al (2016) ask:
  - "Many scientists have said that, 'By 2100, the average U.S. temperature will rise at least 6 degrees Fahrenheit"' and asked them "How many degrees Fahrenheit do you personally expect the average U.S. temperature to rise by 2100, if further regulatory steps are not taken?" Participants could indicate their answer by selecting a number from 0 to 12.
- Not suprisingly, responses were somewhat centered around 6 degrees
- Here anchoring made sense because they were trying to study belief updating by climate change optimists (deniers?) vs pessimists (believers?)
- However, it also shows that anchoring can be dangerous if one is after absolute rather than relative responses.

Framing

- Make sure everybody is thinking about the same object and trade-offs
  - "Do you think we should spent more public funds on health care?"
    - ★ Where does the extra dollar come from? Raising taxes? Military budget?
- How much do you think you will earn in 10 years?
  - With or without inflation?

Measurement error

- If something is very important, ask conceptually similar question multiple times, also in reverse manner.
  - "Do you think you will have enough money to go to university?"
  - "Do you think you will struggle financially at university?"
  - Afterwards can, for instance, extract first factor using principal component analysis.
- Example:

Question with indirect scales
To what extent do you agree or disagree that the income distribution in the US is fair?
- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree

Construct-specific questions
How fair or unfair do you think the distribution of income is in the US?
- ☐ Very unfair
- ☐ Somewhat unfair
- ☐ Neither fair nor unfair
- ☐ Somewhat fair
- ☐ Very fair

- Adapt visual format to type of answer you need (e.g., if you are looking for a dollar amount, put a $ next to the box.)
- If you want an integer, make sure only integers can be used

Question ordering

Guiding principles (which can work in opposite directions)

1. Respondents are often more engaged and less tired at beginning
2. Questions that come at beginning may influence later responses
3. Respondents form opinion about your survey at beginning and it's critical to capture their interest

- If you ask filter questions, (e.g. "Are you self-employed?"), ask all of these first and then ask the follow-ups (e.g. "What is your income from self-employment?")
- Sensitive questions should come later in the survey (religion, age)

Legal and trust issues

- Get ethical approval from your host institution
  - This may (or may not) take much longer than you think. At some institutions committee will only meet once per term.
- Provide contact information, consider IRB requirements, take into account rules such as the GDPR in the EU, and reassure respondents about complete anonymity and confidentiality.
- Provide limited information about the purpose of the study without revealing too much (priming!)

Quality assurance

- If appropriate, inform respondents that careless answers may be flagged and their pay may be withheld.
- Embed attention check:
  - We would now like to ask you a question about the following problem. In surveys like this, it occasionally happens that participants quickly click through the survey without carefully reading the questions. This reduces the quality of the data and impairs the results of research studies. To demonstrate that you read the survey carefully, please answer both "Very interested" and "Not interested at all" to the following question. Considering the above, how interested are you in politics? [Very interested/Reasonably interested/Somewhat interested/Not very interested/Not interested at all]
- Check for identical responses across multiple submissions.
- Delete respondents that are too fast.
- Use CAPTCHA

How to Deal with Bots in Online Surveys

- **Employ Honeypot Fields:**
  - Use hidden fields that should not be filled by human respondents.
  - Flag and filter out responses where these fields are filled.
- **Use ReCAPTCHA and Bot Detection Services:**
  - Integrate advanced bot detection tools (e.g., Google's ReCAPTCHA).

Types and sources of bias

- Types of bias
  - moderacy bias
  - extreme bias
  - ordering bias
  - acquiescence bias
  - experimenter demand effect
  - social desirability bias
- Sources of bias
  - the respondents' behavior (e.g., carelessness or social desirability bias)
  - the content of the question (e.g., leading questions)
  - the design of the questionnaire (e.g., the order of questions that can induce priming)
  - the characteristics of the survey situation itself (e.g., experimenter demand effect)

Biases in answer selection: Moderacy, extreme response, and response-order bias

- Moderacy response bias: the tendency to respond to each question by choosing a category in the middle of the scale.
- Extreme response bias: the tendency to respond with extreme values on the rating scale.
- Response-order bias: when the order of response options in a list or a rating scale influences the response chosen.
  - primacy effect: when respondents are more likely to select one of the first alternatives provided (more common in written surveys).
  - recency effect: occurs when respondents choose one of the last items presented to them (more common in face-to-face or orally presented surveys).

Social desirability bias

- Many exciting questions can be "touchy"
  - Respondents desire to avoid embarrassment and project a favorable image to others, which results in them not revealing their actual attitudes.
- Online surveys likely minimize SDB relative to other survey modes because there is no surveyor in front of respondent
- How to address
  - Repeat reminders of anonymity.
  - Use implicit association test (Lee and Tiptoe (2022) develop one for prejudice against low education groups)

Experimenter demand effect

- Anonymity. Online surveys can relieve social pressure.
- Monetary incentives and real stakes questions
- Obfuscated follow-ups are follow-up studies with the same respondents and dependent variables, but without the respondent knowing the original and follow-up are related.
- Obfuscated information treatments try to obscure the purpose of the experiment, e.g., by giving extra info irrelevant to goal, asking questions about unrelated issues, and giving people an unrelated reason why they receive information.
- Design and question wording
- Hiding the purpose of experiment or study
- Measuring beliefs about the study purpose

Conducting surveys in developing countries

- Pre-negotiate with enumerators. Otherwise you might end up in a hold up trap.
- Be considerate of local culture.
  - You might require permission from chief before holding survey.
- Check patterns in responses carefully. Indications of cheating are:
  - Your names might show up as respondents.
  - Identical responses across surveys.
- Bring enough pens, pencils, etc!
- If you walk around with enumerators, respondents might become very chatty and want to impress.
- Consider (numerical) literacy levels
  - There are methods of eliciting numerical beliefs by using number of days out seven as an example (Sunday might be fix point for church) or giving ten peanuts representing ten percentage points each (Delavande et al 2022 Handbook of Economic Expectations, Delavande et al 2011)

Hypothetical scenarios

Hypothetical scenarios

- Traditionally economists assumed perfect information about production technologies. Moreover, returns are often assumed to be identical for everyone.
  - This is questionable.
- Using hypothetical scenarios you can create your own mini laboratory to study beliefs about different things.
  - You may also want to change a product and wonder what this will do to demand.
- Imagine you want to study how people perceive the return to some investment on earnings.
  - You might also be interested in whether perceived returns vary by sex.
  - And initial level of skills.

You can create your own simulated randomized control trial about beliefs

- How much will person X earn in two years?
  - Imagine person X is a woman and does a low investment.
  - Imagine person X is a woman and does a high investment.
  - Imagine person X is a man and does a low investment.
  - Imagine person X is a man and does a high investment.
- Now you have 4 scenarios per respondent related to investment $k \in L, H$ and sex $j \in W, M$

	Woman	Man
Low investment	Y_L^W	Y_L^M
High investment	Y_H^W	Y_H^M

Your mini beliefs RCT - Fixed effects regression

	Woman	Man
Low investment	Y_L^W	Y_L^M
High investment	Y_H^W	Y_H^M

- You can back out many aspects of perceived returns!

Run fixed effects regression:

$$Y_i^j = \alpha_i + \beta I_H + \gamma W + \delta I_H \times W + \varepsilon$$

- β tells you about perceived return to high investments
- γ tells you about perceived level difference between men and women
- δ tells you about perceived difference in return to high investment for women vs men

Your mini beliefs RCT - One regression for each respondent

Run separate regression for each respondent:

$$Y_k^j = \alpha_i + \beta_i I_H + \gamma_i W + \delta_i I_H \times W + \varepsilon$$

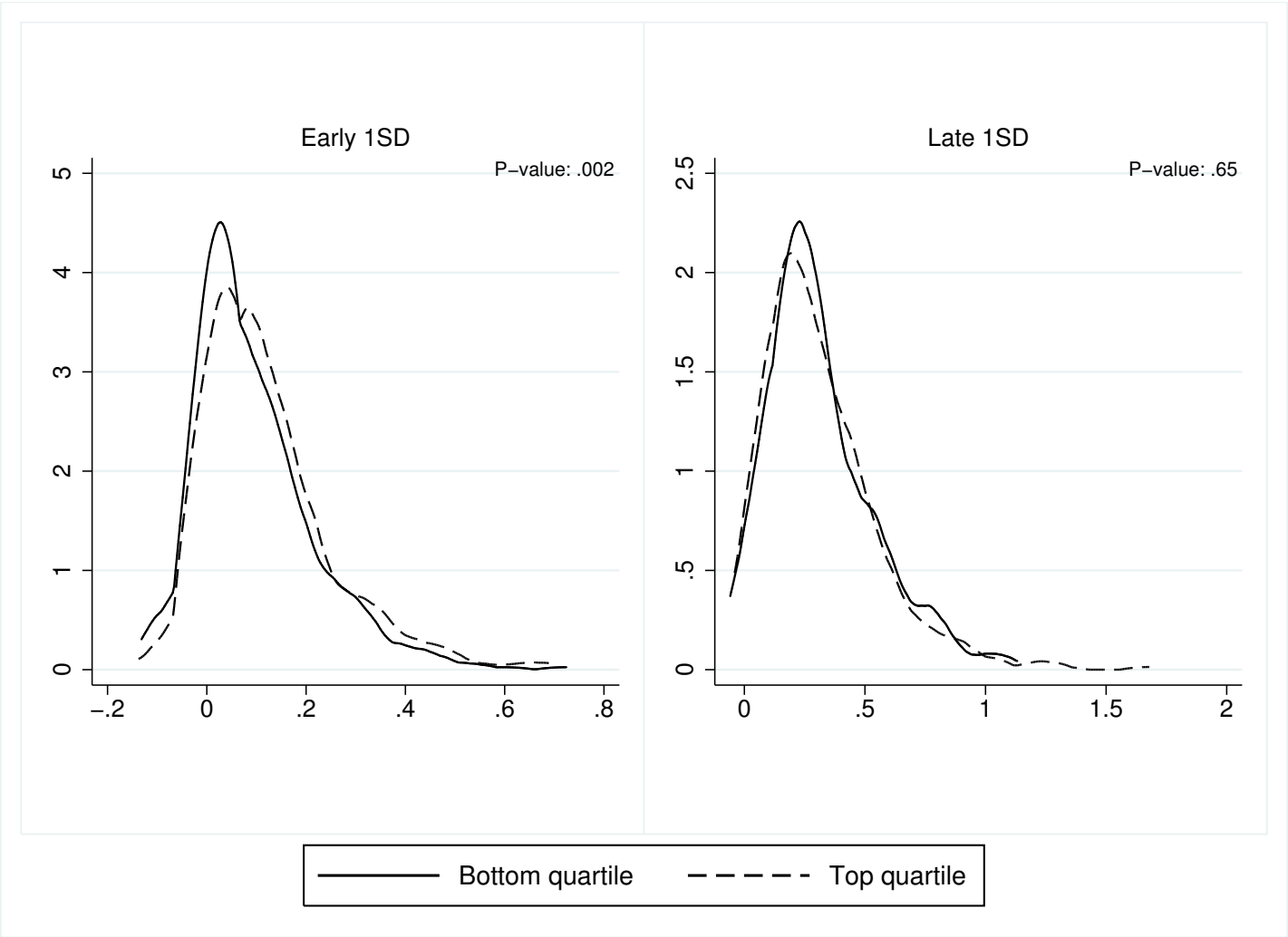
Now, for each respondent you have :

- β_i perceived return to high investments
- γ_i perceived level difference between men and women
- δ_i perceived difference in return to high investment between men and women

and can plot distributions, e.g., by sex of respondent. Alternatives:

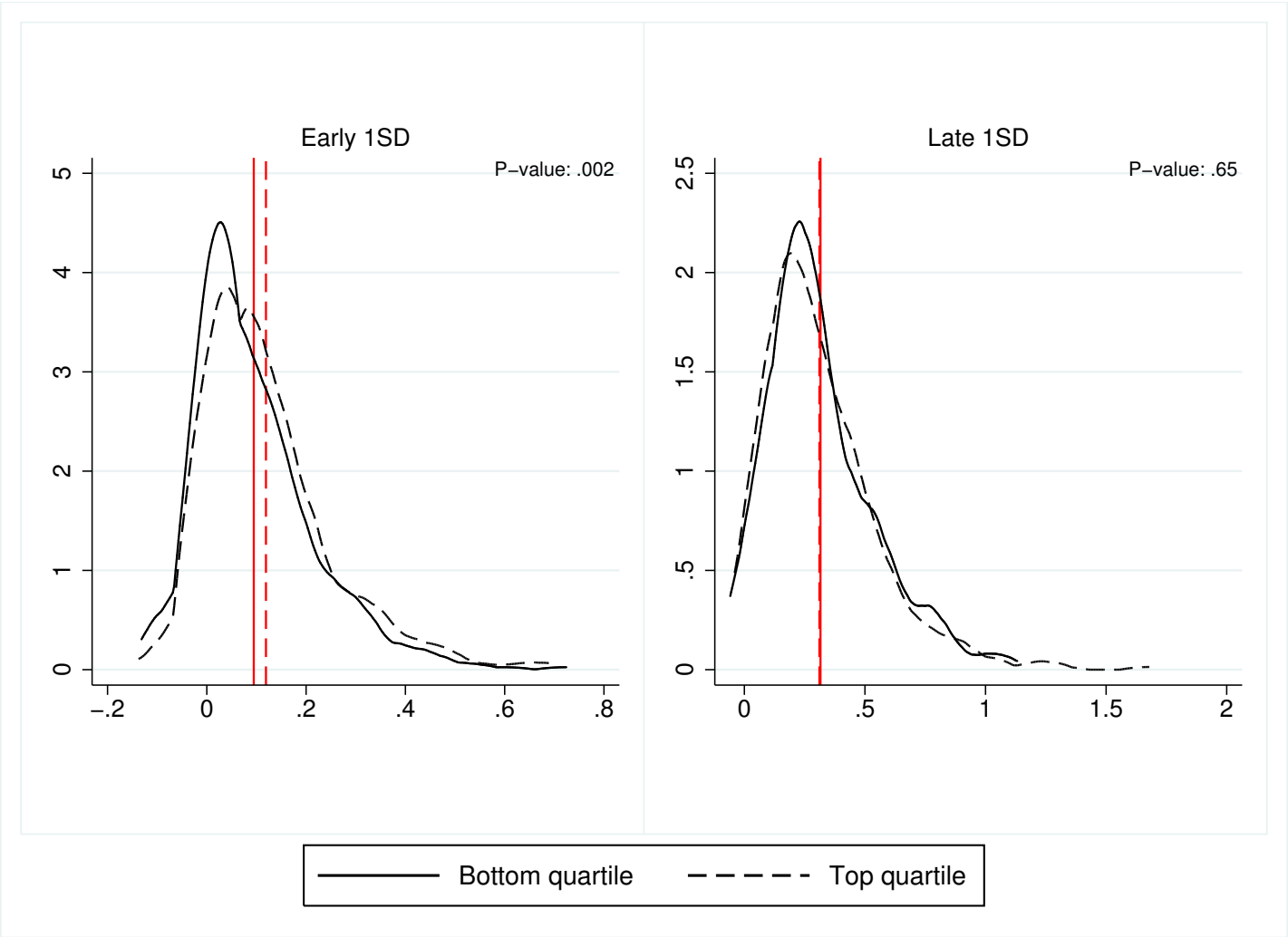
- Run fixed effects regression on full sample and interact with respondent characteristics.
- Calculate differences non-parametrically for each respondent which might make more sense in this example with only 4 scenarios per respondent.

Examples – How much do parents think that investing into their children will effect children's future income



Parents in top income quartile think returns are higher to early investments.

Examples – How much do parents think that investing into their children will effect children's future income



Parents in top income quartile think returns are higher to early investments.

Making scenarios "easy"

- Scenarios are a complex thought experiment
- Prepare respondents for this exercise with an introductory text
  - "We will now ask you a range of hypothetical questions where we ask you to guess the outcome."
  - "These questions are difficult and we want you to think carefully about them."
  - "There are no right or wrong answers. We are interested in what you believe is the outcome."

Who should be the subject of scenarios?

- Scenarios can be constructed wrt to respondent or hypothetical person
  - Returns for self are more pertinent for own decision but return wrt hypothetical person allows keeping everything else equal across respondents.
  - In order for respondents to consider everything else equal across scenarios one has to emphasize this carefully.
    - ★ E.g., "Now imagine Pedro who lives in the same neighborhood, has the same level education and job, but decides to do the high investment."
  - For hypothetical person one can vary gender, age, or other characteristics (which would be difficult for self).

- My advice: If the decision and conditions apply to everyone, then ask about self (e.g., sample of 17 year olds deciding about university). If the scenario is more abstract for individuals then use hypothetical person (e.g., imaging you had a child).

- Reality check: The referee will always demand the one you did not choose.

Vignettes: Step method

- Example: Ask respondents whether they would accept a pay cut for a certain benefit (e.g., being allowed to work from home one day a week)
- Two different ways of doing this
  1. Vary pay cut across respondents
     - ★ Some respondents see 1%, others 5%, others 10%
  2. Start with a given value, e.g., 5%
     - ★ Then depending on whether respondent accepts or rejects one decreases or increases value until they switch.
     - ★ Starting value can be varied across respondents to see whether it changes preferences.
- This method is more appropriate for eliciting preferences over attributes rather than beliefs about returns.

"Real" experiments

Rise in information experiment

Haaland, Roth, and Wohlfart: Designing Information Provision Experiments

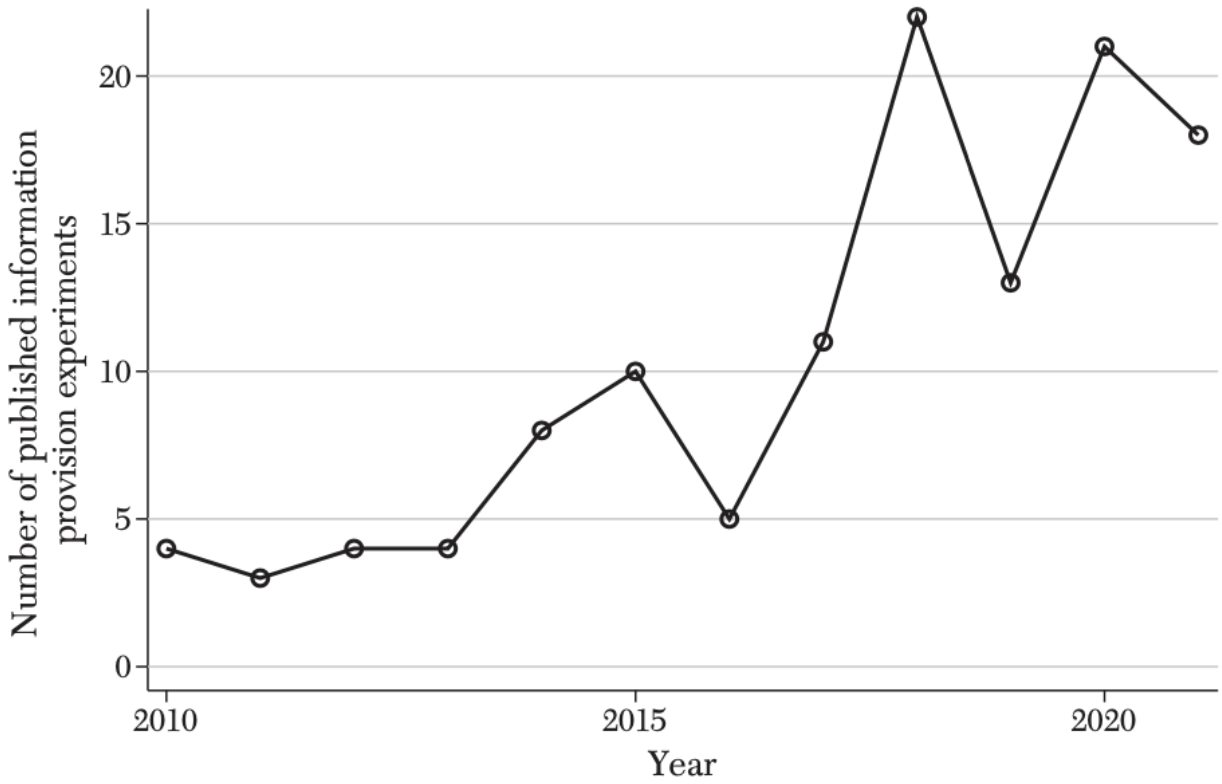


Figure 1. Number of Information Provision Experiments Published in Leading Journals Since 2010

Design choices

- Will you use a between- or within-respondent design?
  - Between-subject: each respondent is only subject to one experimental condition
  - Within-subject: each respondent is subject to multiple experimental conditions. Difference is order in which conditions are administered.
- When will you measure your dependent variable: before, after, or both before and after treatment?
- In information experiments (and sometimes others), different variables to consider:
  - "first stage beliefs" (the belief or information that your treatment is trying to shift)
  - "second stage" (dependent variables influenced by first-stage ones)

Designing the Information Intervention

Key Aspects of Designing Information Interventions

- **Types of Information Provided:**
  - *Quantitative Information:* Statistics, expert forecasts, and data-driven insights.
  - *Anecdotal Evidence and Narratives:* Qualitative information through stories, videos, or case studies.
  - *Tailored vs. General Information:* Information specific to individuals versus broader general data.
- **Sources of Information:**
  - Official statistics, expert opinions, and historical data.
  - Qualitative stories or media influences.
- **Presentation of Information:**
  - Consider how to present data to minimize cognitive load and improve comprehension.
- **Credibly Identifying Effects:**
  - Distinguishing the effects of information from mere priming effects.

Cross-Learning in Information Provision Experiments

Cross-learning occurs when respondents update beliefs about variables other than the intended object of interest after receiving new information.

- Coibion, Georgarakos et al. (2023): Information about inflation not only altered respondents' inflation expectations but also their beliefs about GDP growth.

Challenges:

- Makes it harder to isolate the effect of information on behavior via belief changes.

Mitigating Cross-Learning Effects

- Fixing Beliefs About Other Variables:
  - Provide identical information about other variables to both control and treatment groups.
- Trade-offs:
  - Multiple pieces of information may dilute attention and weaken the first stage of the experiment.

Persistence of effects of information provision

- It is sometimes possible to resurvey respondents
  - Make sure to discuss this with survey company before launching first wave!
  - Costs are higher
  - Response rates drop exponentially over time
- Very useful if one wants to demonstrate that treatment effect is real and not due experimenter demand effect
  - Example: Treatment shows that inequality is very high in your country. After that respondents say inequality is bad. Will they still exhibit treatment effect one month later?

Obfuscated follow-up surveys

- Re-surveying respondents while pretending that this is a completely different survey
  - Different company, color, topic, etc
- Goal: Overcome experimenter demand effect
  - If obfuscation is well done respondent will not realize they are being surveyed about the same topic
- Problem
  - Hard to "hide" question related to information treatment
  - If one finds no effect, is that because it did persist or because now one is asking a different question?

Typical seminar/referee questions

- "Did you randomize the order?"
  - Sometimes asking one question before another might prime respondents into certain behavior
    - ★ "Do you think it is important to invest time into children?"
    - ★ "How much time do you invest into your child"
  - Randomize the order and don't forget to keep track of the randomization!
- "Do you weight your results?"
  - Add robustness results to Appendix where you "correct" for non-response or biased sample.

Using monetary incentives and real stakes questions

- Monetary incentives for truthful revelation can be used to reward accurate answers or particular behaviors.
- Real stakes questions can be used to lend credibility to self-reported attitudes, behaviors, and beliefs
  - Can include e.g., petitions or donations to charities related to the issue of interest
- Making respondents "spectators" can help respondents internalize their self-reported choices. As spectators, they observe the actions or choices of other respondents (the "stakeholders") and then allocate rewards to them. Spectators' choices are then implemented with a certain probability.
  - "You will be matched with another participant who will receive X if you do Y."

Incentivizing correct responses

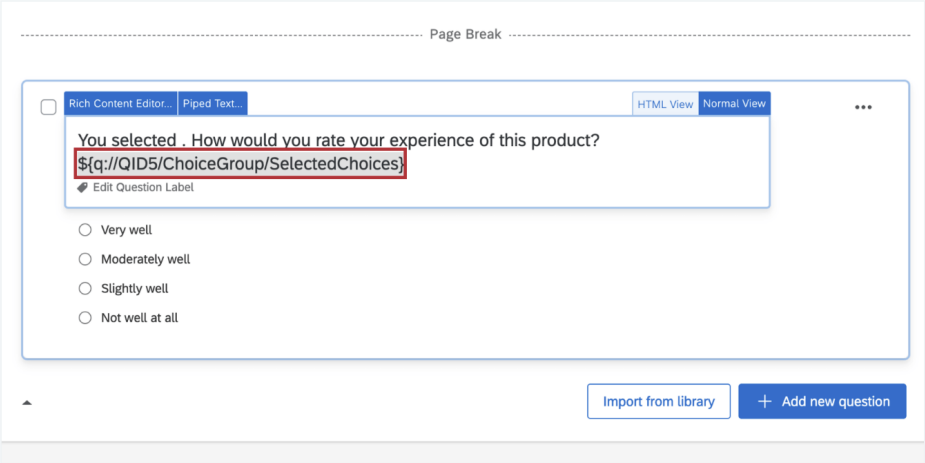
- Many experiments elicit guesses about certain observable statistics or distributions (earnings, growth, inflation, ...)
- "Incentives, Search Engines, and the Elicitation of Subjective Beliefs: Evidence from Representative Online Survey Experiments" by Grewenig et al (2021)
  - Incentivize belief accuracy affects stated beliefs about average earnings by professional degree and average public school spending.
  - Incentive provision does not impact earnings beliefs, but improves school-spending beliefs.
  - Response patterns suggest that the latter effect likely reflects increased online-search activity.
  - Consistently, an experiment that just encourages search-engine usage produces very similar results.

Steps in Programming a Qualtrics Survey

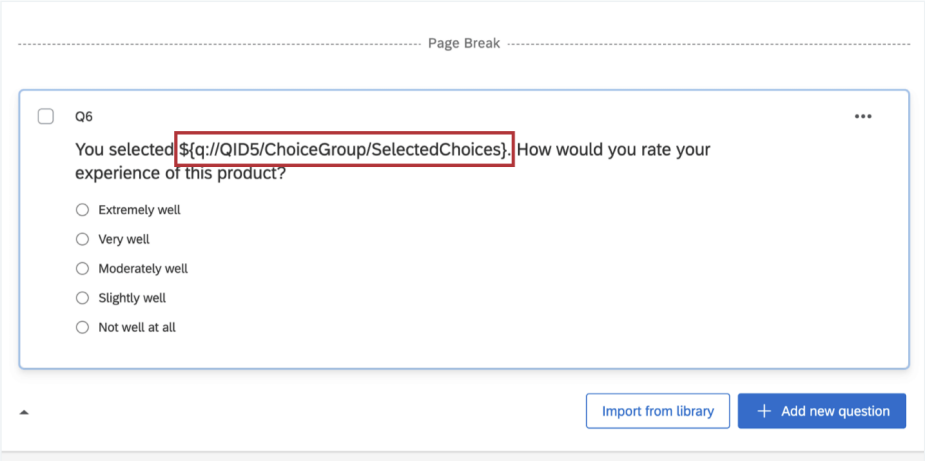
1. **Create a New Survey**
   - Start by selecting "Create a Project" in the Qualtrics dashboard.
2. **Design the Survey Structure**
   - Outline the flow of your survey, including different blocks and questions.
   - Use "Survey Flow" to organize blocks and add logic (e.g., randomization, branching).
3. **Add Questions**
   - Use the "Question Builder" to add various types of questions (e.g., multiple choice, text entry, Likert scale).
   - Customize question settings, such as validation, randomization, and display logic.
4. **Apply Logic and Branching**
   - Set up logic for displaying questions based on previous responses.
   - Use branching to direct respondents to different parts of the survey based on their answers.

Programming a survey - piped text example

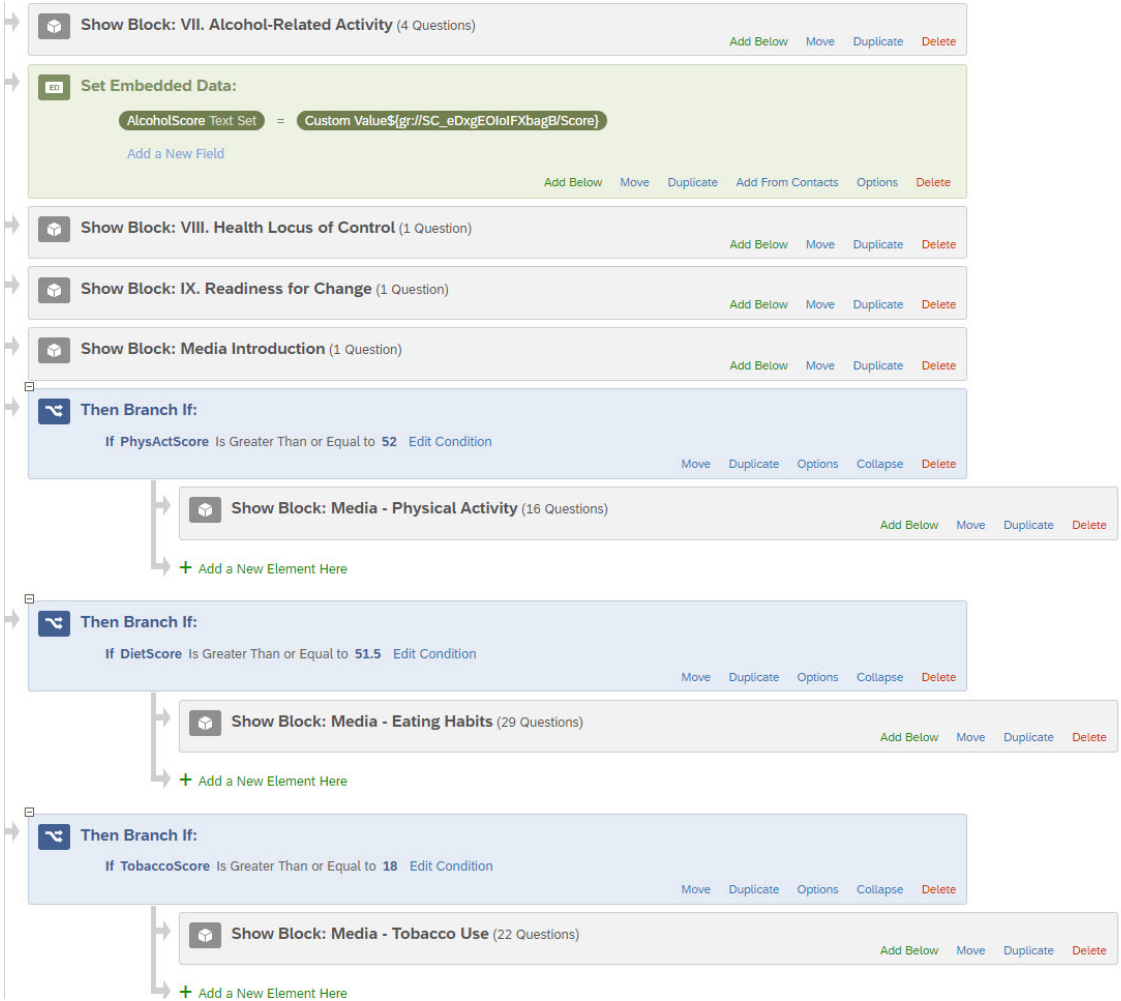
2 Notice that the piped text is inserted into your survey, as shown by a code surrounded by a dollar sign and brackets.



3 Place the piped text where you want the actual text to appear in the question.



Programming a survey - flow example



Steps in Programming a Qualtrics Survey

5. Test and Preview the Survey

- Use the "Preview" mode to test the survey for flow, logic, and functionality.
- Conduct a soft launch to identify any issues before full deployment.

6. Distribute the Survey

- Generate an anonymous link, distribute via email, or integrate with panels.
- Monitor response rates and adjust distribution methods as needed.

7. Analyze and Export Data

- Use Qualtrics' built-in tools to analyze responses.
- Export data to CSV, SPSS, or other formats for further analysis.

When you think you are done

"Simulation" – The final rehearsal

- Code analysis in Stata/R/Python
- Try survey multiple times yourself, and ask friends, family, and colleagues to have a go
- Generate many descriptives
  - Check if randomizations worked and are recorded!
- Conduct analysis as if it were real

And again

- Tighten survey
  - Delete questions not required
- Pilot survey yourself and with contacts again
- Now you might be ready to pilot the survey using the company

Carefully monitor the entire survey process

- It is useful to have a contact person at survey company. Not all offer this.
  - We have repeatedly worked with same contact which makes process smooth (and can get you a discount).
- Monitor data collection in real-time and adapt to unforeseen circumstances.
  - Are dropout rates very high? This might mean that the survey is too convoluted or there is a bug somewhere.
  - Monitor quotas. Survey company might not be targeting the right people.
- Run your (pre-)analysis codes developed in rehearsal stage early in the process to spot inconsistencies.
  - Resist temptation of stopping when results are "good".

Costs

- Bargain
- ALWAYS program the survey yourself
  - This not only makes a gigantic difference for costs but also puts you in control (and likely requires less work)
  - However, survey flows can be complicated processes with dead ends and unwanted jumps. Don't underestimate.
- One efficient way of running a mini pilot with a big sample is by seeding/testing a question for your next project in your current project.

Summary

- Technology has made surveys cheaper and better
- Think carefully about how to design your survey and what exactly you are after.
- Keep it short and simple (but be creative).
- Test, test, test....

Research examples

Beliefs about maternal labor supply
2024
by Boneva, Golin, Kaufmann, and Rauh

Survey design

Belief elicitation: Perceived return to maternal labor supply

- Design two sets of vignettes (randomized across respondents)
  - Vig. A: Elicit perceived total effect by varying labor supply (and income) of the mother while her child is 1-5 years old
  - Vig. B: Elicit perceived income effect by varying household income only
- Vignettes A and B feature an average German married couple
  - Spouses are both 30 yo and have a one-year-old child
  - Before birth of child both worked FT and each earned 38K gross/year
  - Mother takes 12 months maternity leave while father keeps working FT
  - Mother wants to return to FT work after maternity leave
  - Do not want to have additional children
  - Household expenditure decisions taken jointly

Survey design

Belief elicitation: Perceived return to maternal labor supply

- Design two sets of vignettes (randomized across respondents)
  - Vig. A: Elicit perceived total effect by varying labor supply (and income) of the mother while her child is 1-5 years old
  - Vig. B: Elicit perceived income effect by varying household income only
- Vignettes A and B feature an average German married couple
  - Spouses are both 30 yo and have a one-year-old child
  - Before birth of child both worked FT and each earned 38K gross/year
  - Mother takes 12 months maternity leave while father keeps working FT
  - Mother wants to return to FT work after maternity leave
  - Do not want to have additional children
  - Household expenditure decisions taken jointly

Survey design

Belief elicitation: Perceived return to maternal labor supply

	Mother	Father	Household Income
Vignettes A			
Scenario 1	Stays home (0k)	Works full-time (40k)	40k gross/year
Scenario 2	Works part-time (20k)	Works full-time (40k)	60k gross/year
Scenario 3	Works full-time (40k)	Works full-time (40k)	80k gross/year
Vignettes B			
Scenario 1	Stays home (0k)	Works full-time (40k)	40k gross/year
Scenario 2	Stays home (0k)	Works full-time (60k)	60k gross/year
Scenario 3	Stays home (0k)	Works full-time (80k)	80k gross/year

- Vignettes A: Decided by chance whether the family gets a place in a childcare center and whether the center is open for the full day
- Vignettes B: Decided by chance whether the new employer that opens a new office nearby offers the father a better-paid job

Survey design

Belief elicitation: Perceived return to maternal labor supply

	Mother	Father	Household Income
Vignettes A			
Scenario 1	Stays home (0k)	Works full-time (40k)	40k gross/year
Scenario 2	Works part-time (20k)	Works full-time (40k)	60k gross/year
Scenario 3	Works full-time (40k)	Works full-time (40k)	80k gross/year
Vignettes B			
Scenario 1	Stays home (0k)	Works full-time (40k)	40k gross/year
Scenario 2	Stays home (0k)	Works full-time (60k)	60k gross/year
Scenario 3	Stays home (0k)	Works full-time (80k)	80k gross/year

- Vignettes A: Decided by chance whether the family gets a place in a childcare center and whether the center is open for the full day
- Vignettes B: Decided by chance whether the new employer that opens a new office nearby offers the father a better-paid job

Survey design

Belief elicitation: Perceived return to maternal labor supply

	Mother	Father	Household Income
Vignettes A			
Scenario 1	Stays home (0k)	Works full-time (40k)	40k gross/year
Scenario 2	Works part-time (20k)	Works full-time (40k)	60k gross/year
Scenario 3	Works full-time (40k)	Works full-time (40k)	80k gross/year
Vignettes B			
Scenario 1	Stays home (0k)	Works full-time (40k)	40k gross/year
Scenario 2	Stays home (0k)	Works full-time (60k)	60k gross/year
Scenario 3	Stays home (0k)	Works full-time (80k)	80k gross/year

- Vignettes A: Decided by chance whether the family gets a place in a childcare center and whether the center is open for the full day
- Vignettes B: Decided by chance whether the new employer that opens a new office nearby offers the father a better-paid job

Survey design

Belief elicitation: Perceived return to maternal labor supply

Outcomes	*Scale*
Child outcomes	
Vocabulary	Relative rank
Intelligence	(0-100)
Concentration	
Work independently	
Social skills	
Family outcomes	
Satisfaction child	Relative rank
Satisfaction mother	(0-100)
Satisfaction father	
Mother-child relationship	
Mother-father relationship	
*Maternal labor market outcomes**	
Prob. full-time job (age 36)	Probability (0-100%)
Earnings (age 36)	Euro
Earnings (age 42)	Euro

* Only measured in vig.A

Survey design

Belief elicitation: Childcare, labor supply, social norms

- Beliefs about childcare
  - Likelihood of finding a place in childcare
  - Likelihood that the childcare is open the full day
  - Likelihood that the childcare is of high quality
  - Costs of childcare
- Beliefs about social norms
  - Perceived approval of family
  - Perceived approval of friends
- Intended maternal labor supply when child is 1-5 yo
  - Baseline
  - Policy scenario 1: Full-time childcare available
  - Policy scenario 2: Full-time, high-quality childcare available

Survey design

Belief elicitation: Childcare, labor supply, social norms

- Beliefs about childcare
  - Likelihood of finding a place in childcare
  - Likelihood that the childcare is open the full day
  - Likelihood that the childcare is of high quality
  - Costs of childcare
- Beliefs about social norms
  - Perceived approval of family
  - Perceived approval of friends
- Intended maternal labor supply when child is 1-5 yo
  - Baseline
  - Policy scenario 1: Full-time childcare available
  - Policy scenario 2: Full-time, high-quality childcare available

Survey design

Belief elicitation: Childcare, labor supply, social norms

- Beliefs about childcare
  - Likelihood of finding a place in childcare
  - Likelihood that the childcare is open the full day
  - Likelihood that the childcare is of high quality
  - Costs of childcare
- Beliefs about social norms
  - Perceived approval of family
  - Perceived approval of friends
- Intended maternal labor supply when child is 1-5 yo
  - Baseline
  - Policy scenario 1: Full-time childcare available
  - Policy scenario 2: Full-time, high-quality childcare available

Data collection

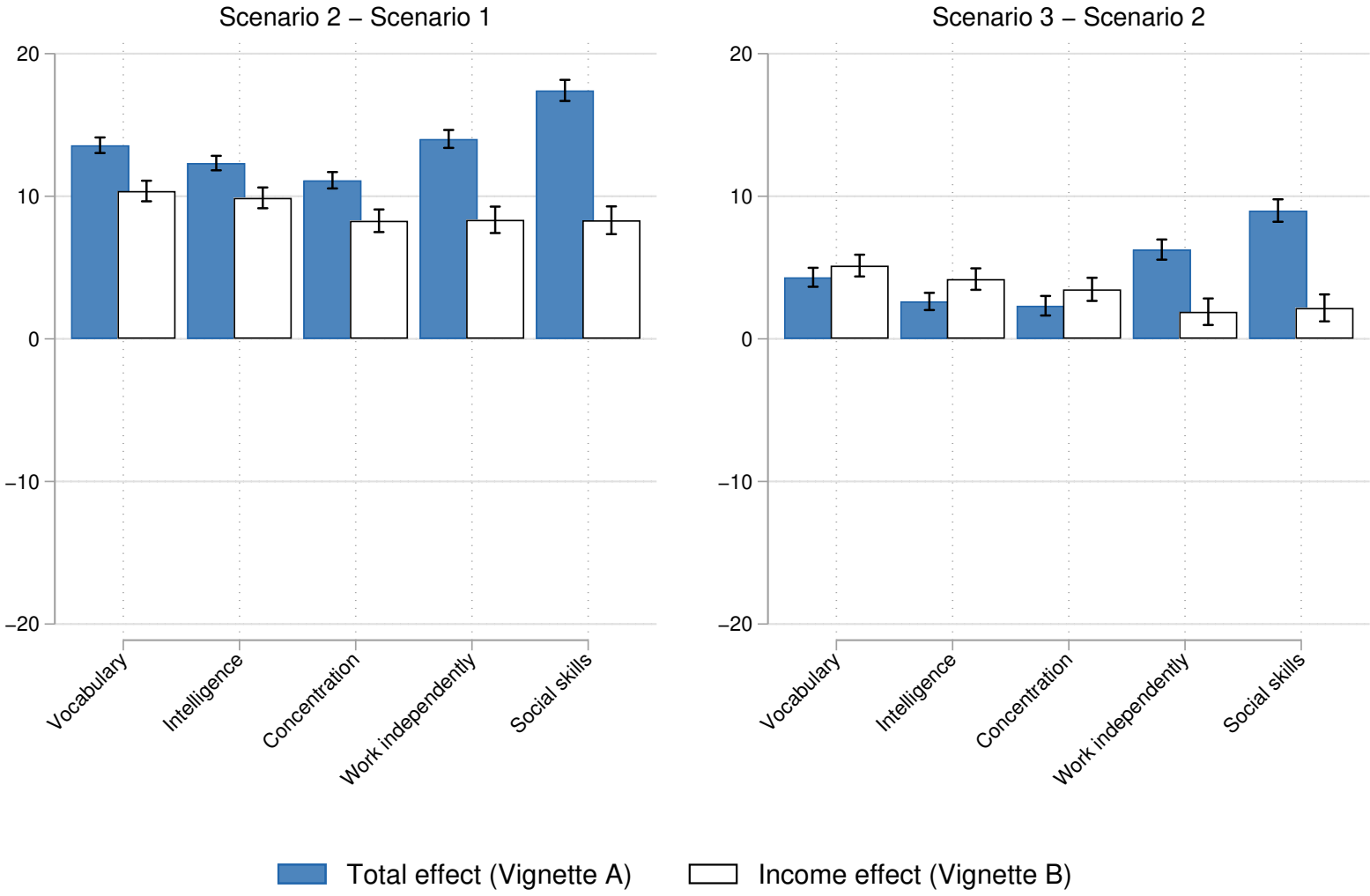
- survey data from a representative sample of 4,000 German adults (aged 18-45) without children
- Data collected in March-May 2022 in collaboration with *Pureprofile*
- Quota-based sampling to ensure representativeness over:
  - Gender
  - Broad educational attainment
  - State of residence

Sample representativeness

	Sample	National population
Female	0.44	0.43
University degree	0.22	0.26
Age*	29.55	28.64
Married*	0.15	0.15
Migrant background*	0.24	0.29
States		
Baden-Wuerttemberg	0.13	0.14
Bayern	0.15	0.15
Berlin	0.05	0.05
Brandenburg	0.02	0.02
Bremen	0.01	0.01
Hamburg	0.02	0.02
Hessen	0.08	0.09
Mecklenburg-Vorpommern	0.02	0.02
Niedersachsen	0.09	0.09
Nordrhein-Westfalen	0.24	0.24
Rheinland-Pfalz	0.04	0.04
Saarland	0.01	0.01
Sachsen	0.05	0.05
Sachsen-Anhalt	0.03	0.03
Schleswig-Holstein	0.03	0.03
Thueringen	0.02	0.02

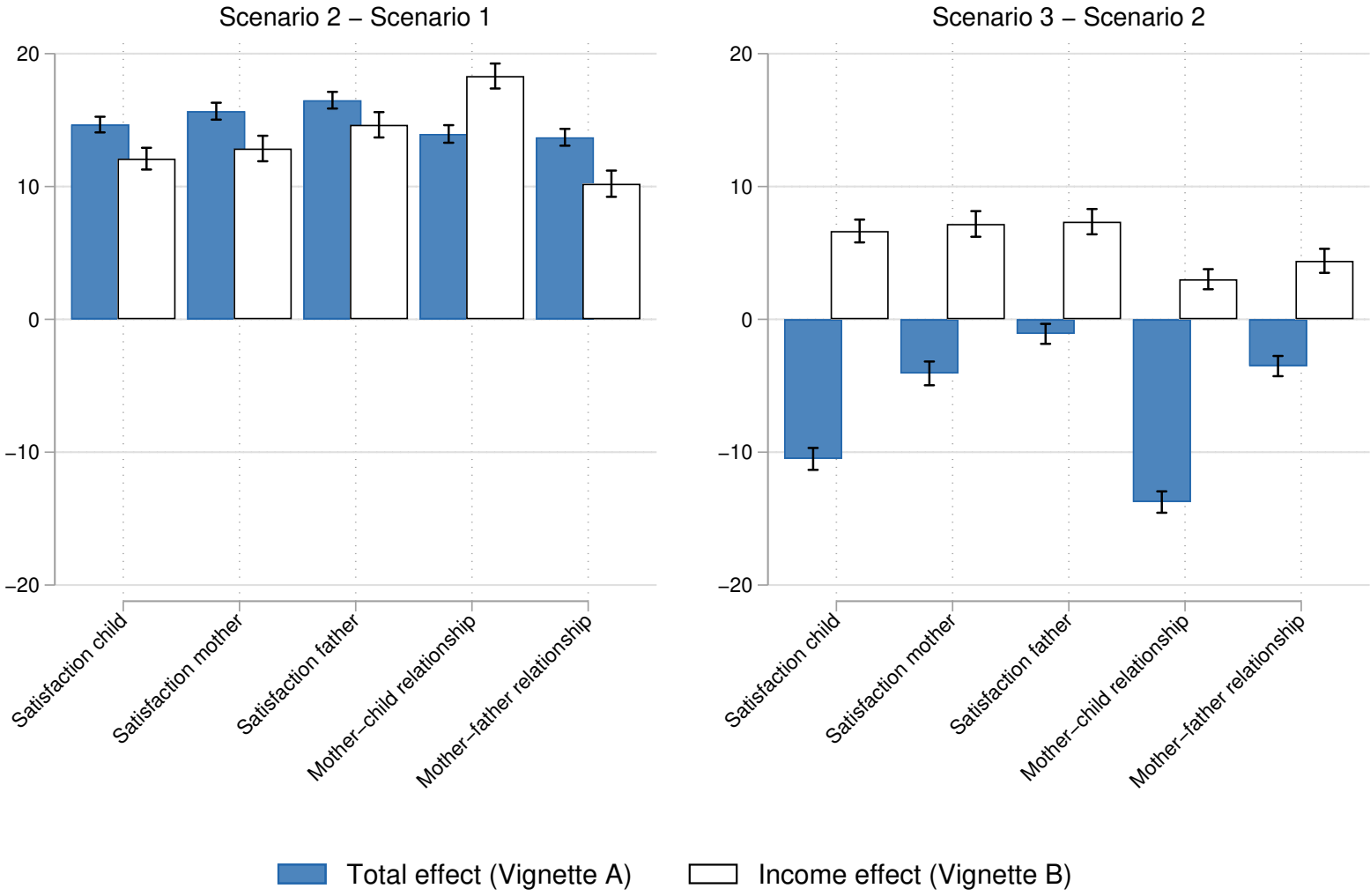
Evidence on beliefs about returns

Child outcomes



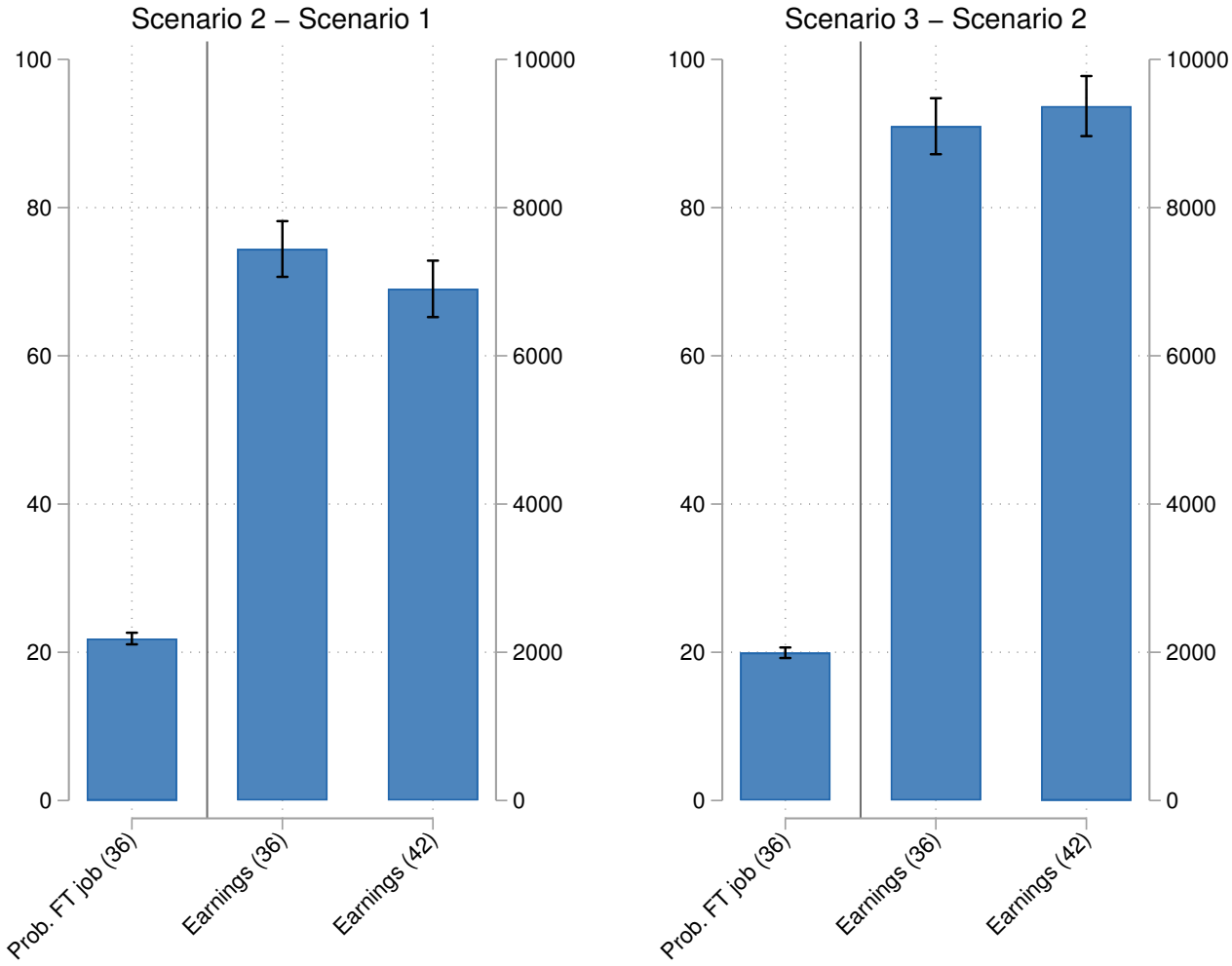
Evidence on beliefs about returns

Family outcomes



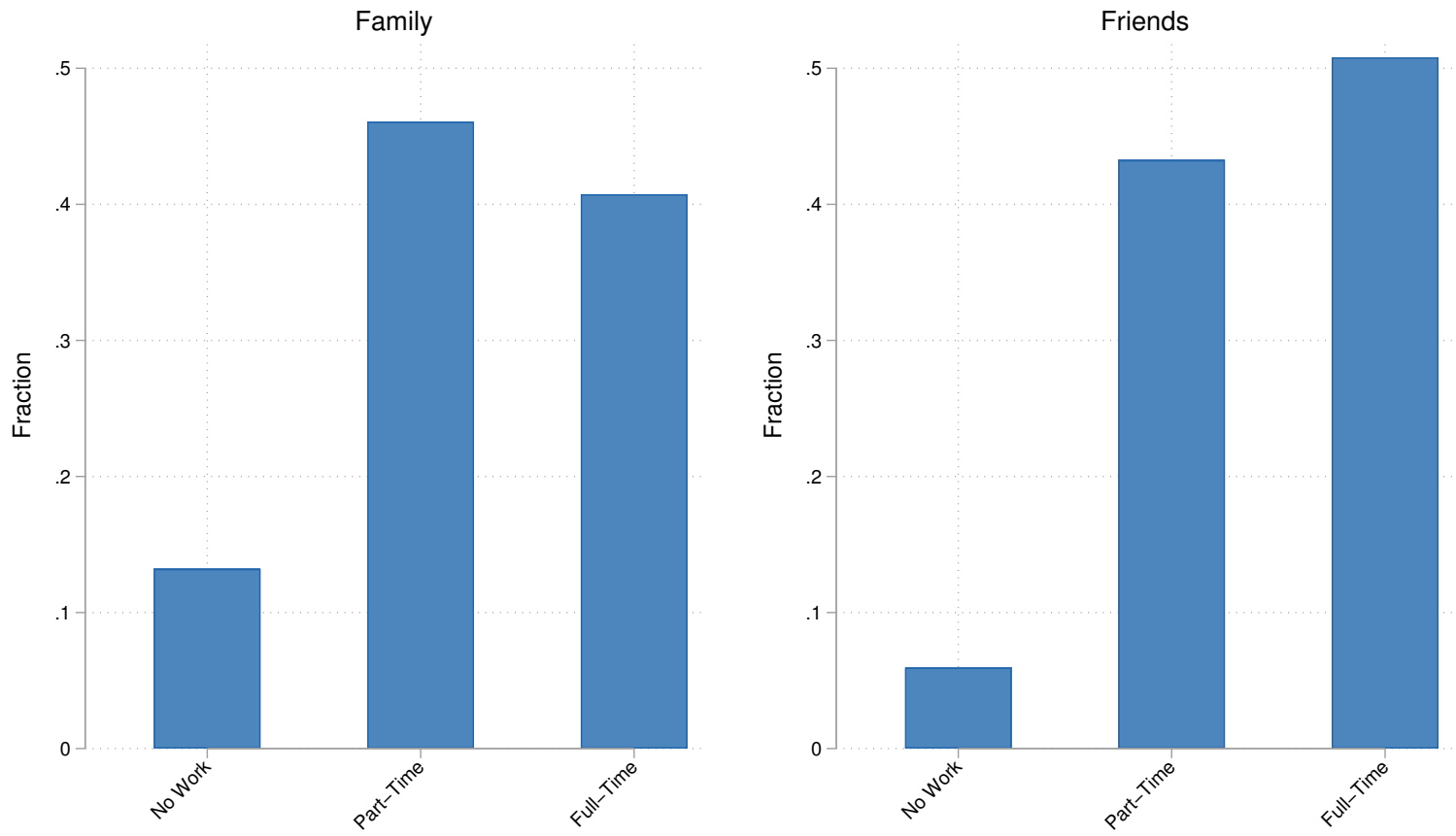
Evidence on beliefs about returns

Maternal labor market outcomes

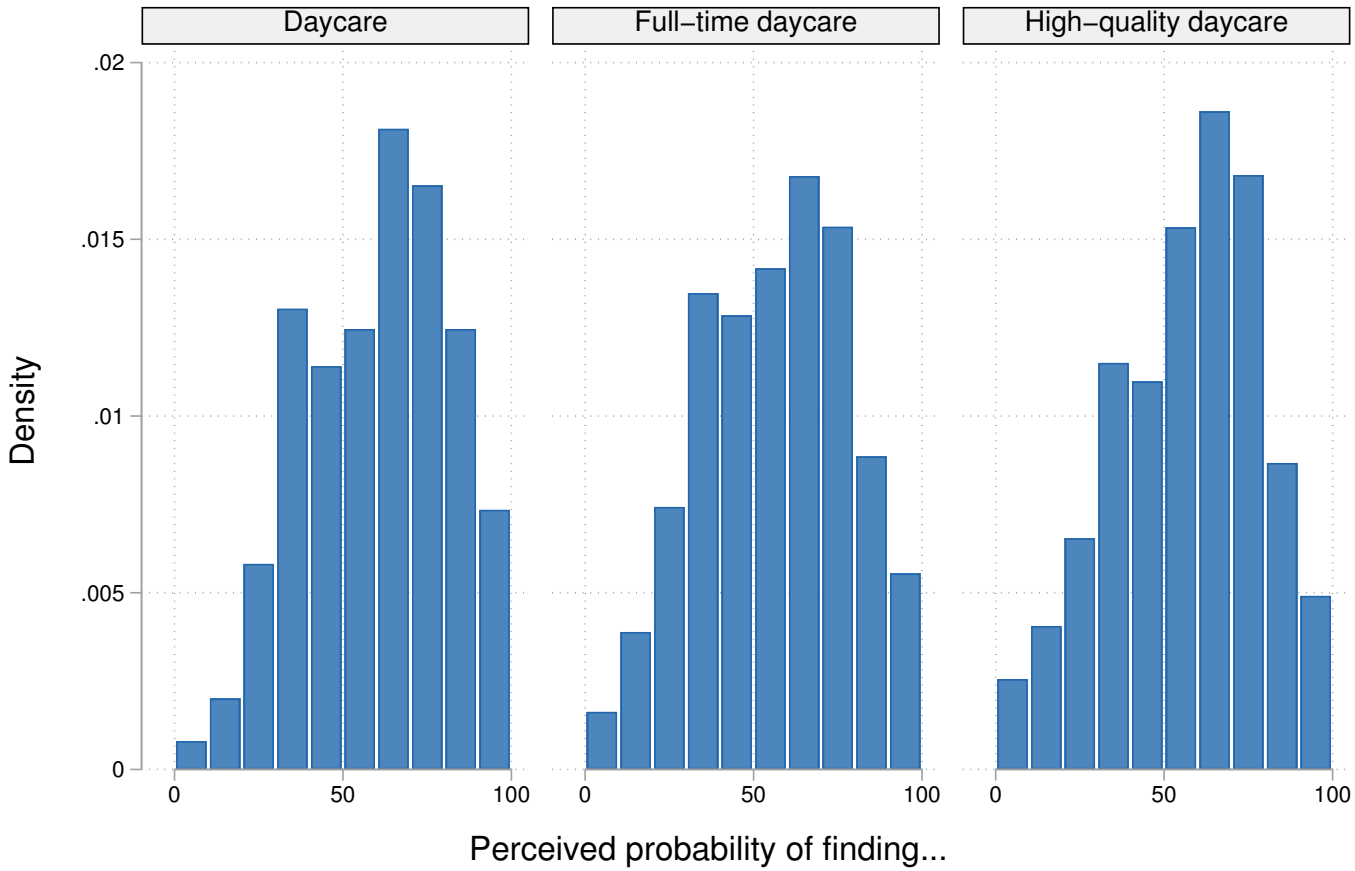


Evidence on beliefs about social norms

'Imagine you have a child and a full-day place in childcare was available to you. What do you think the following people would approve of most? That you [your partner/the mother of your child] are [is]...'

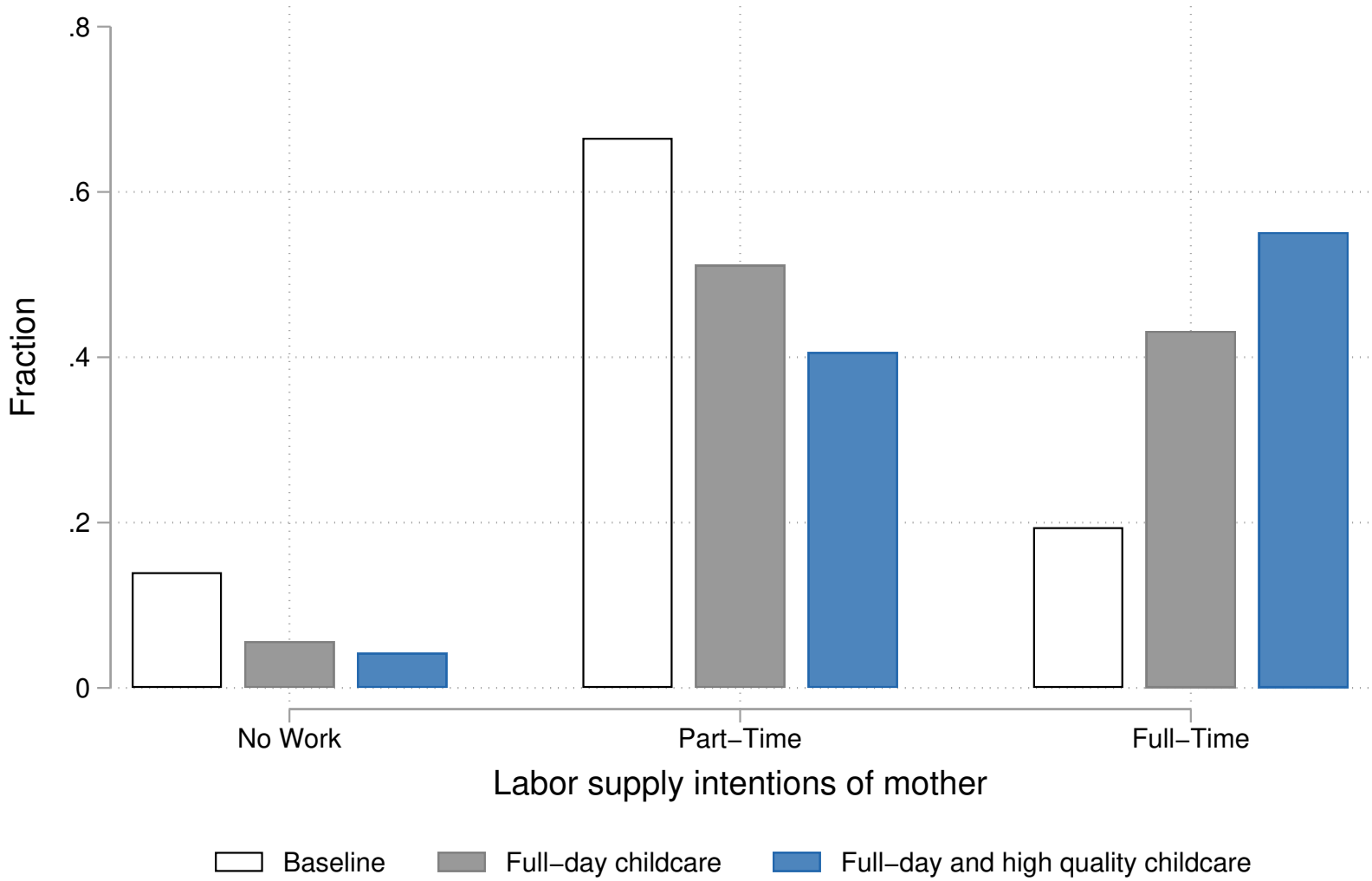


Evidence on beliefs about childcare



Average perceived chance of finding childcare (full-time childcare) [high-quality childcare]: 58% (54%) [55%]

Evidence on labor supply intentions



Do beliefs predict labor supply intentions?

- Estimate multinomial probit choice model where individual i can decide (i) not to work, (ii) to work part-time, or (iii) to work full-time when the child is young

$$u_{ij} = \alpha_j + \beta_1 h^C_{2ij} + \beta_2 h^F_{2ij} + \gamma \rho_{ij} Y^m_{2ij} + \delta s_{1ij} + \lambda_j p_i + \xi_j X_i + \varepsilon_{ij}.$$

  - Individual i selects alternative j to maximize utility derived from their choice, u_{ij}
  - The probility that alternative j is chosen is then $Pr(i \text{ chooses } j) = Pr(u_{ik} \leq u_{ij}) \forall k \neq j$

- Estimate three separate choice models for choices made in baseline scenario, and under the policy scenarios where constraints about childcare availability and quality are relaxed

Beliefs strongly predict labor supply intentions

		Childcare	
	Baseline	Full-time	Full-time & high quality
Child skills	0.2990	0.6398***	1.2615***
	(0.2265)	(0.2317)	(0.2794)
Family outcomes	1.8938***	1.6438***	1.4621***
	(0.2651)	(0.2649)	(0.2832)
Maternal earnings (36) - in 000's Euro	-0.0016	0.0047**	0.0126***
	(0.0020)	(0.0021)	(0.0029)
Family's opinion	0.3377***	0.3799***	0.4276***
	(0.0474)	(0.0484)	(0.0586)
Friends' opinion	0.1842***	0.3700***	0.3809***
	(0.0491)	(0.0546)	(0.0615)
Observations	2873	2873	2873
Controls	Yes	Yes	Yes

Marginal effects - Alternative-specific variables

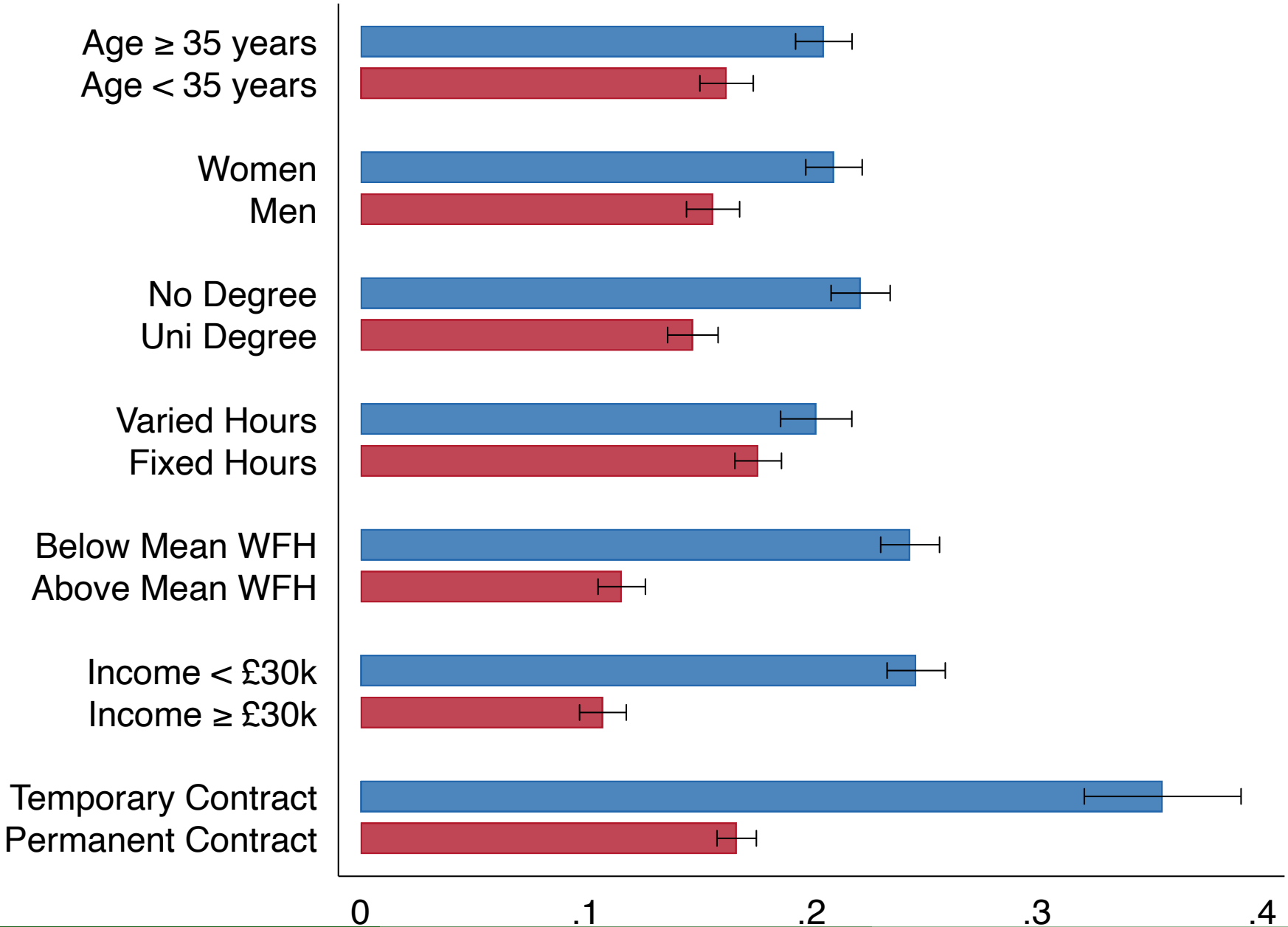
Full-time and high quality

Heterogeneity in beliefs

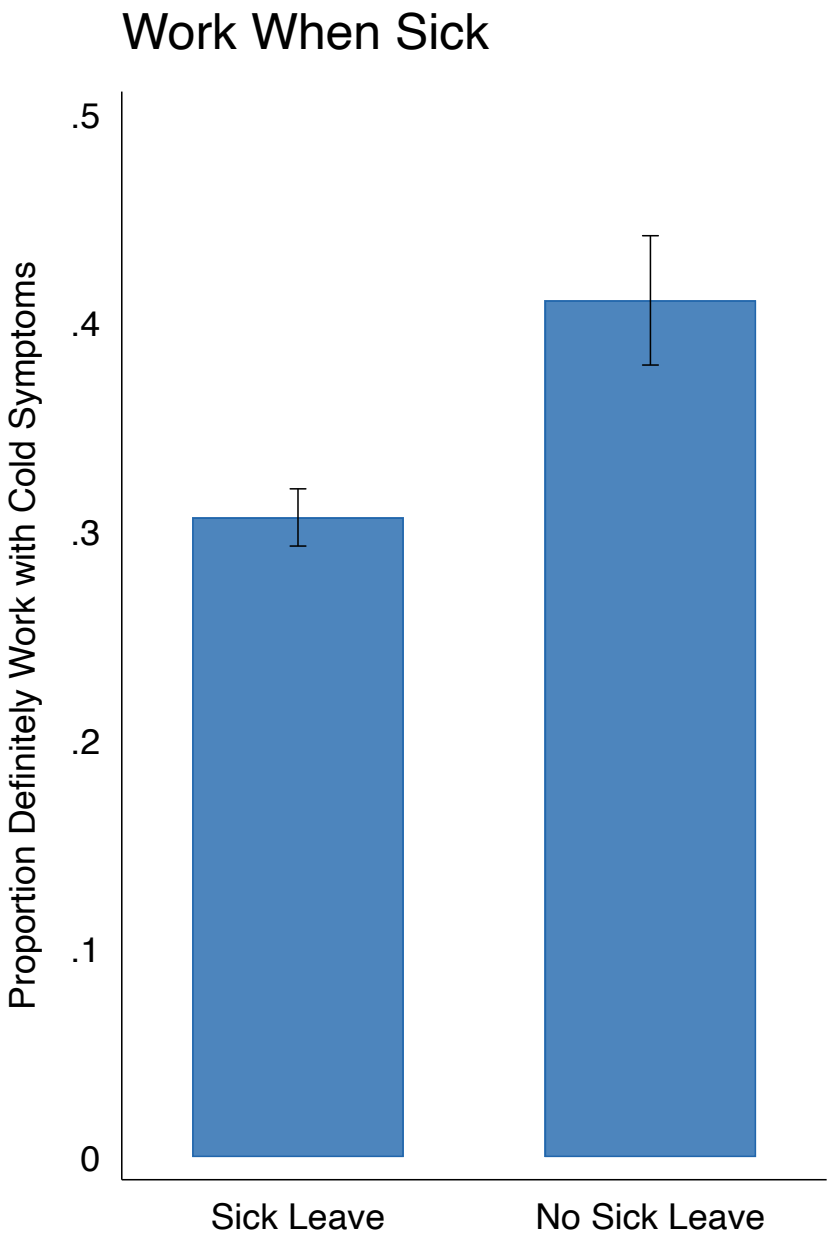
- Substantial heterogeneity in beliefs about returns
- Systematic differences in beliefs across groups in the population
  - Respondents whose mother worked (either part-time or full-time) when they were young perceive the returns to mothers working full-time as significantly higher
  - Perceived returns are higher for respondents who grew up in East Germany

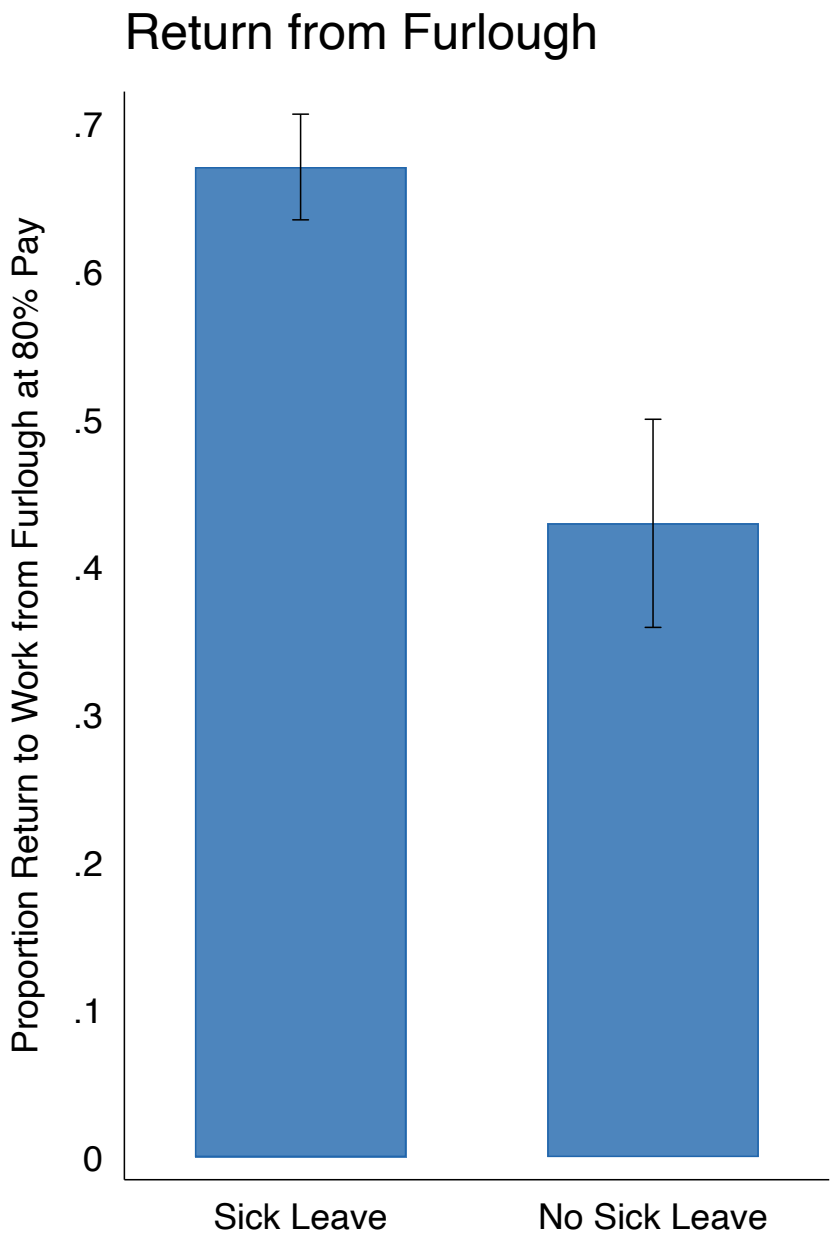
The Value of Sick Pay
European Economic Review (2022)
Adams-Prassl, Boneva, Golin, and Rauh

No access to sick pay by characteristics



Workers without sick pay more likely to work when sick





Vignettes – Hypothetical choice between two contracts

- Employees who are currently not entitled to paid sick leave (WTP):

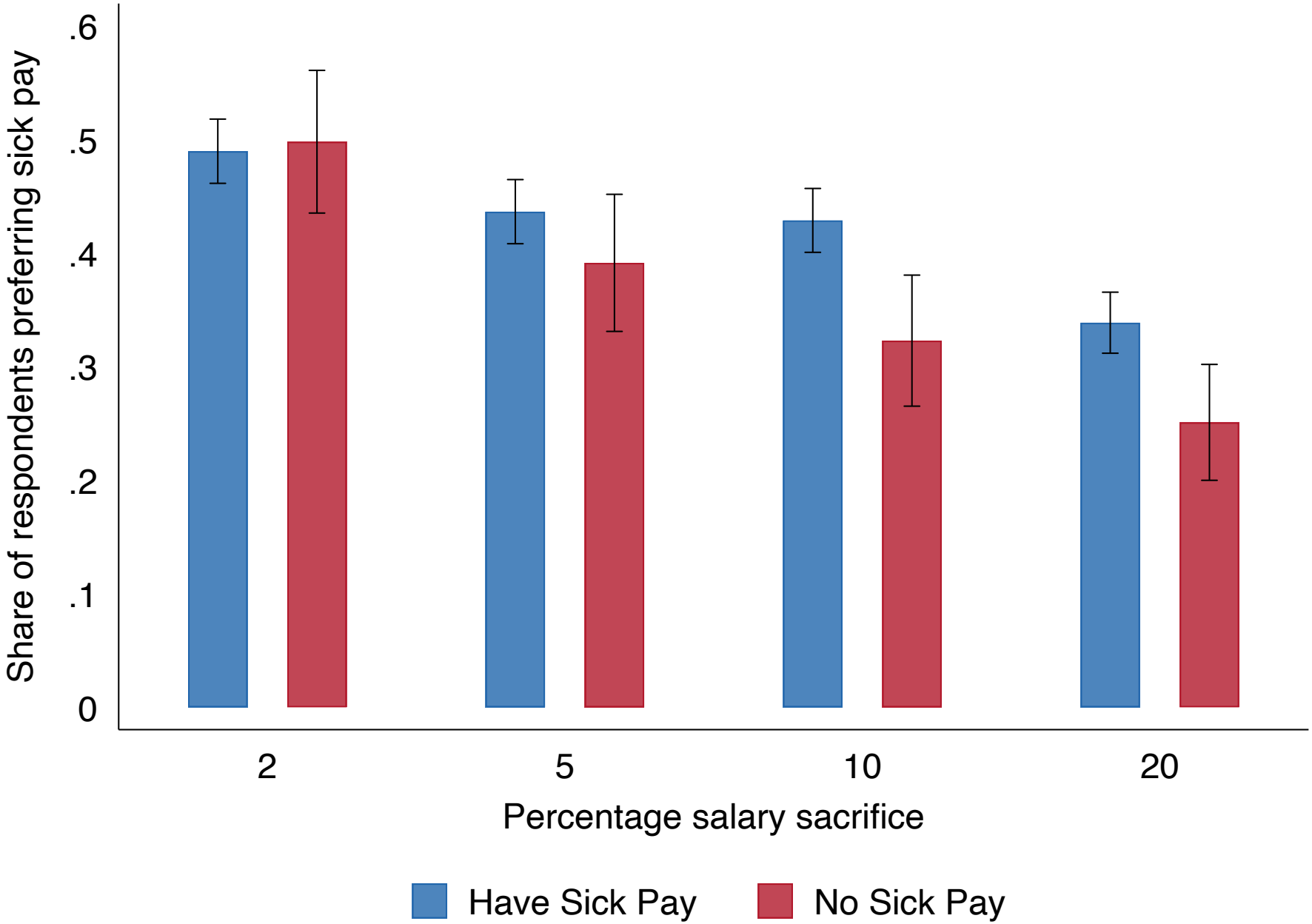
  Suppose your employer in your main job offers you 14 days of paid sick leave per year (in addition to statutory sick pay). In exchange for having access to sick pay you would get X% lower pay per hour. All other aspects of your job would stay the same. Would you accept this arrangement if given the choice? [Yes, No]

- Employees who are currently entitled to paid sick leave (WTA):

  Suppose your employer in your main job offers you X% higher pay per hour. In exchange for having higher pay you would lose your entitlement to paid sick leave through your employer (hence, you would only have access to statutory sick pay). All other aspects of your job would stay the same. Would you accept this arrangement if given the choice? [Yes, No]

- 'X' randomly varies across respondents (2, 5, 10 or 20).

Access to sick pay and willingness to pay



Individual WTP and WTA

	(1)	(2)	(3)	(4)
No Paid Sick Leave	-0.0975***	-0.0973***	-0.1011***	-0.1008***
	(0.0182)	(0.0182)	(0.0183)	(0.0183)
5% Salary Increase / Decrease	-0.0640***	-0.0641***	-0.0664***	-0.0664***
	(0.0177)	(0.0177)	(0.0177)	(0.0177)
10% Salary Increase / Decrease	-0.0824***	-0.0831***	-0.0842***	-0.0848***
	(0.0177)	(0.0177)	(0.0176)	(0.0177)
20% Salary Increase / Decrease	-0.1719***	-0.1727***	-0.1770***	-0.1779***
	(0.0174)	(0.0174)	(0.0173)	(0.0174)
Observations	5725	5717	5744	5736
R^2	0.1059	0.1058	0.1137	0.1136
Controls	yes	yes	yes	yes
Region F.E.	yes	yes	yes	yes
Wave F.E.	yes	yes	yes	yes
Occupation F.E.	no	no	yes	yes

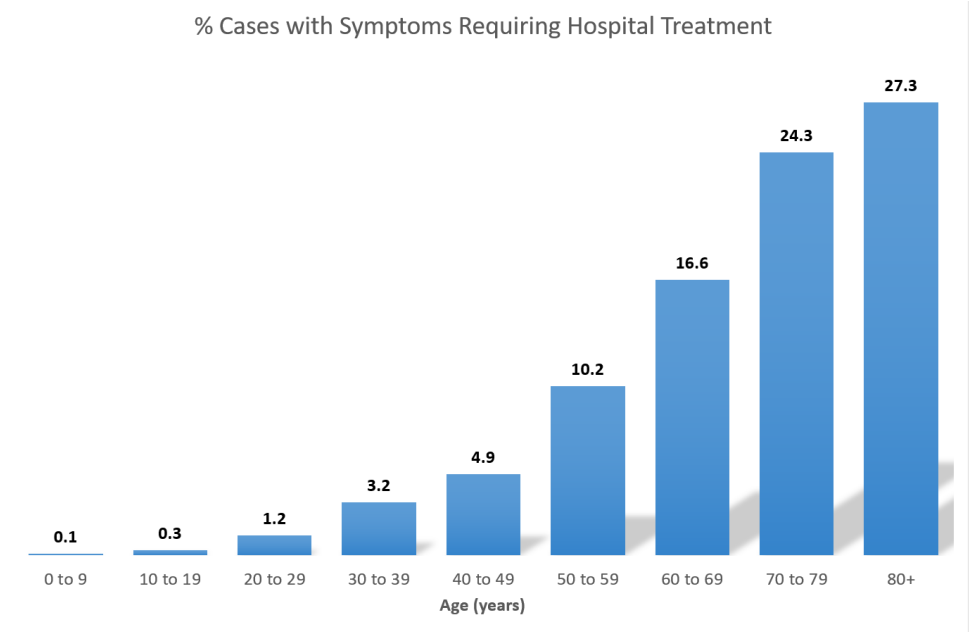
Treatments

1. Truthful information on health-related aspects of the coronavirus outbreak
2. Health information + economic consequences of the outbreak.
3. Health information + information on workers without paid sick leave.

Each group forms 25% of sample.

Health information

In different countries, officials predict that more than 70% of people might get infected. While most people will only develop mild symptoms, the virus can be severe for older people, many of whom may require hospital treatment. This has already put a lot of pressure on the health systems in countries where the outbreak started earlier. For each age group, the chart below shows the estimated proportion of coronavirus cases with symptoms that need hospital treatment.



Economic information

The virus is predicted to have a big impact on the whole economy. In the UK, economists predict that around 700,000 people will lose their jobs during the crisis. In the United States, unemployment has already risen sharply. 281,000 people became unemployed in the week ending 14 March, a sharp rise from 211,000 in the previous week. This rise is larger than any week-to-week unemployment movement during (or since) the 2008 financial crisis. Many businesses have already been affected by a fall in revenue caused by social-distancing measures. Anne**, a small business owner, has seen many orders cancelled. With no cash coming in, she says she was forced to lay off most of her 17 employees.*

** Estimates as of 24 March 2020 from KPMG and Capital Economics.*

*** Not her real name*

March 15 2020, The New York Times

Information on others and externalities

Many people are not entitled to paid sick leave. This puts them in a difficult situation if they risk losing their job or their income if they stay home. Adam, who is self-employed, said in response to the outbreak: "If you're self-employed you have to continue working. I'm not about to make my children starve because of coronavirus. If I'm physically able to work, then isolation is not happening for me." Adam admits that continuing to work might spread the virus. "That's a risk I would have to take", he said. Not granting paid sick leave to all workers poses serious threats to public health.*

**Not his real name.*

March 7 2020, The Guardian.

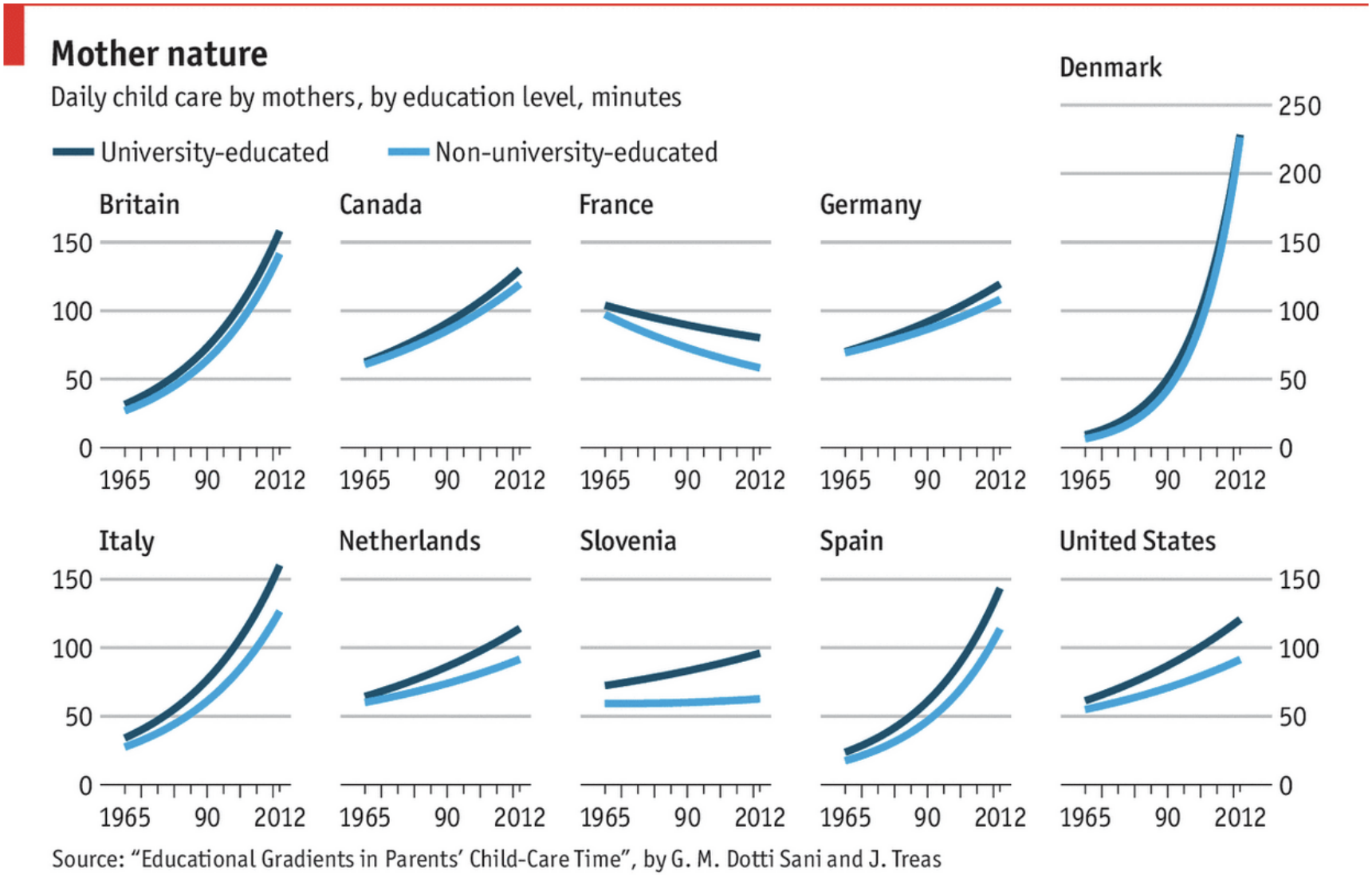
Treatment effects

	Ind. WTP	14 Days Sick Leave			Self-Employed Sick Leave		
		Neutral	Agree	Str. agree	Neutral	Agree	Str. agree
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Health	0.0046	0.0208	0.0174	0.0325	0.0170	0.0116	0.0582**
	(0.0270)	(0.0145)	(0.0235)	(0.0274)	(0.0156)	(0.0245)	(0.0268)
Health + Econ	-0.0148	-0.0005	0.0279	0.0143	0.0144	0.0194	0.0539**
	(0.0264)	(0.0156)	(0.0238)	(0.0278)	(0.0160)	(0.0250)	(0.0272)
Health + Sick Pay	-0.0016	0.0316**	0.0358	0.0380	0.0277*	0.0324	0.0765***
	(0.0265)	(0.0139)	(0.0232)	(0.0276)	(0.0151)	(0.0243)	(0.0270)
Observations	2515	2516	2516	2516	2517	2517	2517
R^2	0.1223	0.0388	0.0547	0.0375	0.0441	0.0588	0.0255
Mean of control group	0.4508	0.9175	0.7524	0.3968	0.9048	0.7206	0.3175
Controls	yes	yes	yes	yes	yes	yes	yes

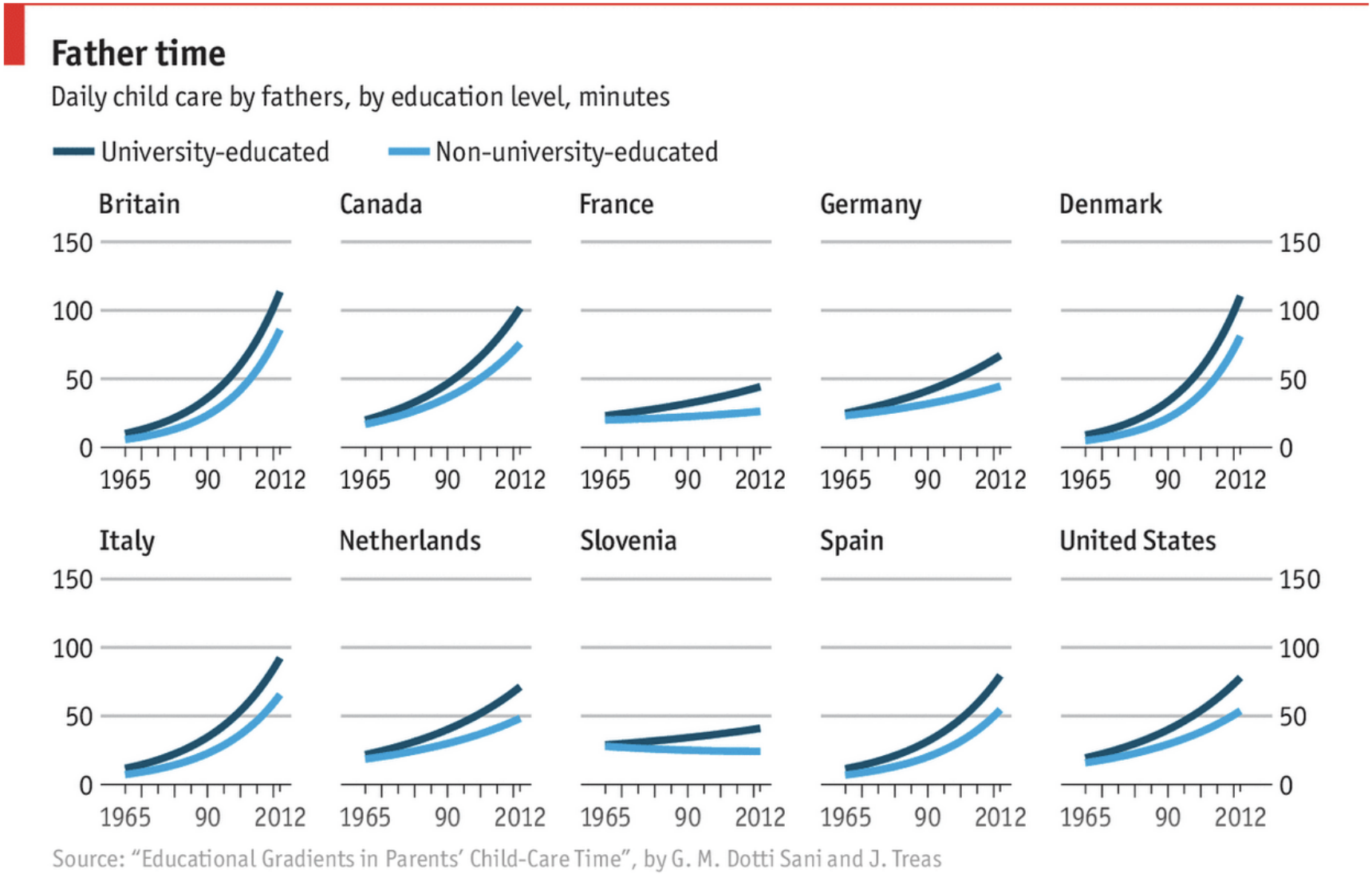
- No effect on individual WTP.
- Increased support for sick pay for self-employed.

Parental beliefs about returns to educational investments – the later the better?
Journal of European Economic Association (2022)
by Boneva and Rauh

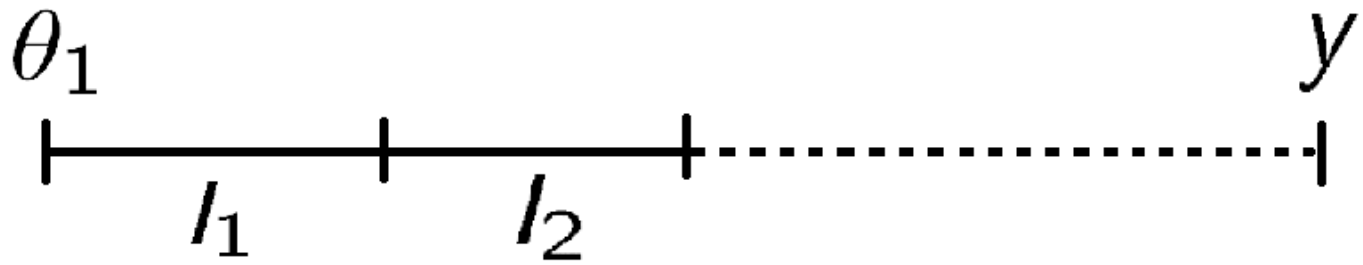
Maternal time investments



Paternal time investments



Theoretical framework



Theoretical framework

- investment choices driven by *perceived* function:

$$y = h_i(\theta_1, I_1, I_2)$$

- perceived marginal returns:

$$\frac{\partial h_i(\cdot)}{\partial I_1}, \frac{\partial h_i(\cdot)}{\partial I_2}$$

- perceived complementarity:

$$\frac{\partial h_i(\cdot)}{\partial I_2 \partial I_1} \gtreqless 0, \frac{\partial h_i(\cdot)}{\partial I_t \partial \theta_1} \gtreqless 0$$

Theoretical framework

- investment choices driven by *perceived* function:

$$y = h_i(\theta_1, I_1, I_2)$$

- perceived marginal returns:

$$\frac{\partial h_i(\cdot)}{\partial I_1}, \frac{\partial h_i(\cdot)}{\partial I_2}$$

- perceived complementarity:

$$\frac{\partial h_i(\cdot)}{\partial I_2 \partial I_1} \gtreqless 0, \frac{\partial h_i(\cdot)}{\partial I_t \partial \theta_1} \gtreqless 0$$

Theoretical framework

- investment choices driven by *perceived* function:

$$y = h_i(\theta_1, I_1, I_2)$$

- perceived marginal returns:

$$\frac{\partial h_i(\cdot)}{\partial I_1}, \frac{\partial h_i(\cdot)}{\partial I_2}$$

- perceived complementarity:

$$\frac{\partial h_i(\cdot)}{\partial I_2 \partial I_1} \lessgtr 0, \frac{\partial h_i(\cdot)}{\partial I_t \partial \theta_1} \lessgtr 0$$

Eliciting Parental Beliefs

Use hypothetical investment scenarios

- vary initial human capital level (θ_1)
- vary level of early investment (I_1)
- vary level of late investment (I_2)

Ask parents about

- expected earnings at age 30

Eliciting Parental Beliefs

Use hypothetical investment scenarios

- vary initial human capital level (θ_1)
- vary level of early investment (I_1)
- vary level of late investment (I_2)

Ask parents about

- expected earnings at age 30

Eliciting parental beliefs

A: The Jones

High Initial Human Capital

	Low Late Investment	High Late Investment
Low Early Investment	(1) Low early/ Low late	(2) Low early/ High late
High Early Investment	(3) High early/ Low late	(4) High early/ High late

B: The Smiths

Low Initial Human Capital

	Low Late Investment	High Late Investment
Low Early Investment	(5) Low early/ Low late	(6) Low early/ High late
High Early Investment	(7) High early/ Low late	(8) High early/ High late

Eliciting parental beliefs

Mr and Mrs Jones have one child, John. John is in Year 3 of primary school, and in the KS1 SATs John achieved the expected level (i.e. Level 2). In the following school years, Mr and Mrs Jones can decide how much to help John with his school work.

Assuming there is no inflation, what do you expect John's gross yearly earnings to be when he is 30 years old if they help John...

... 1 hour every week in school years 3-6, and 1 hour every week in school years 7-10?

Eliciting parental beliefs

Types and Levels of Inputs (Sample A):

- Initial Human Capital (Year 2)
  - Low: did not achieve expected level in Year 2
  - High: did achieve expected level in Year 2
- Early Investments (school years 3-6)
  - Low: spend 1h (or 0h)/week helping child with school work
  - High: spend 4h (or 3h)/week helping child with school work
- Late Investments (school years 7-10)
  - Low: spend 1h (or 0h)/week helping child with school work
  - High: spend 4h (or 3h)/week helping child with school work

Eliciting parental beliefs

Types and Levels of Inputs (Sample A):

- Initial Human Capital (Year 2)
  - Low: did not achieve expected level in Year 2
  - High: did achieve expected level in Year 2
- Early Investments (school years 3-6)
  - Low: spend 1h (or 0h)/week helping child with school work
  - High: spend 4h (or 3h)/week helping child with school work
- Late Investments (school years 7-10)
  - Low: spend 1h (or 0h)/week helping child with school work
  - High: spend 4h (or 3h)/week helping child with school work

Eliciting parental beliefs

Types and Levels of Inputs (Sample A):

- Initial Human Capital (Year 2)
  - Low: did not achieve expected level in Year 2
  - High: did achieve expected level in Year 2
- Early Investments (school years 3-6)
  - Low: spend 1h (or 0h)/week helping child with school work
  - High: spend 4h (or 3h)/week helping child with school work
- Late Investments (school years 7-10)
  - Low: spend 1h (or 0h)/week helping child with school work
  - High: spend 4h (or 3h)/week helping child with school work

Eliciting parental beliefs

Types and Levels of Inputs (Sample B):

- Initial Human Capital (age 5)
  - Low: scored worse than 70% on intelligence test
  - High: scored better than 70% on intelligence test
- Early Investments (age 5)
  - Low: read every 2nd day, rarely go to playground, watch TV 2h/day (-0.5 SDs)
  - High: read every day, go to playground once every fortnight, watch TV 1h/day (+0.5 SDs)
- Late Investments (age 10)
  - Low: moderate interest in child's education, don't talk much to child, sometimes engage in activities together (-0.5 SDs)
  - High: a lot of interest in child's education, talk to child a lot, often engage in activities together (+0.5 SDs)

Eliciting parental beliefs

Types and Levels of Inputs (Sample B):

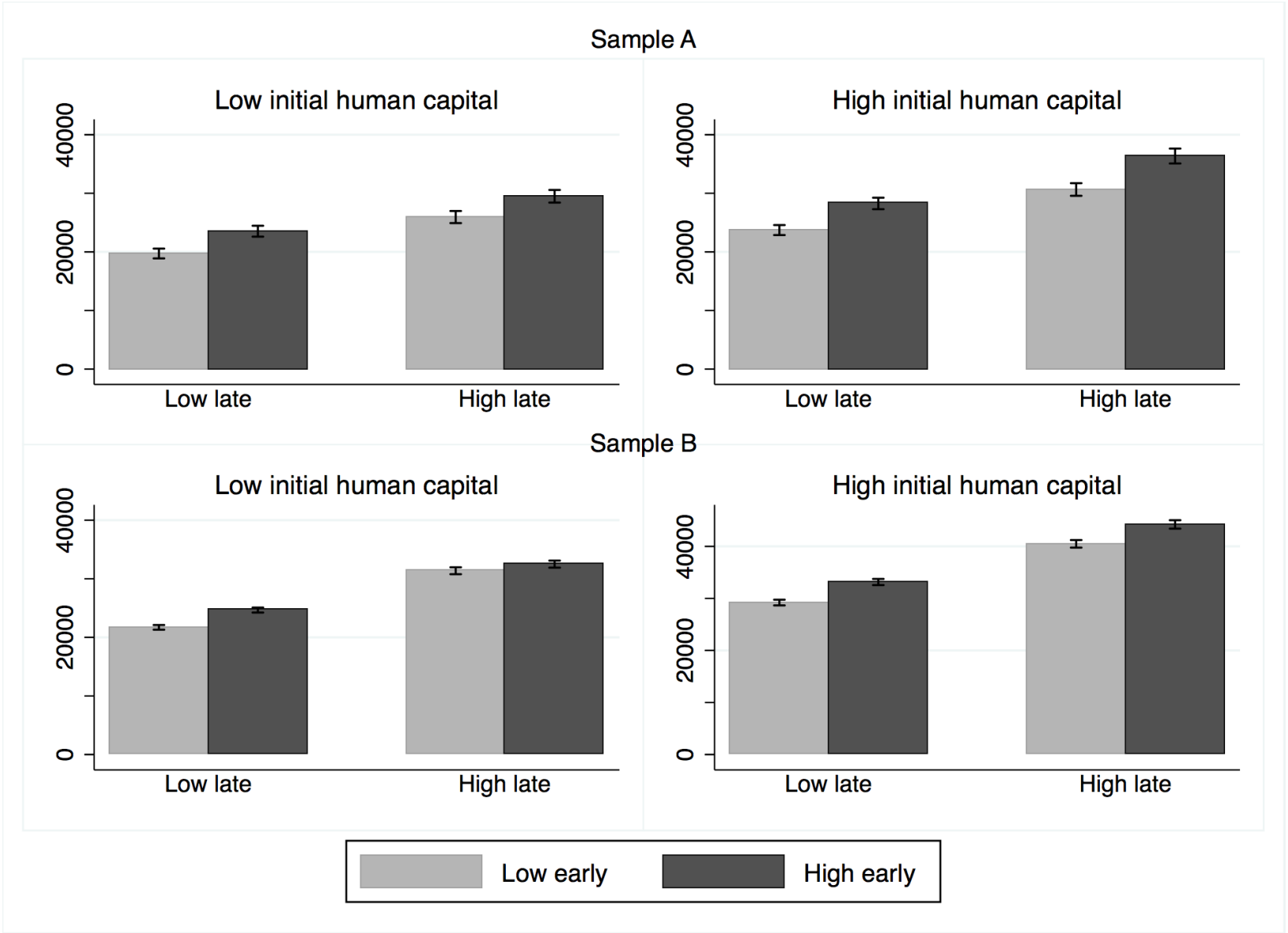
- Initial Human Capital (age 5)
  - Low: scored worse than 70% on intelligence test
  - High: scored better than 70% on intelligence test
- Early Investments (age 5)
  - Low: read every 2nd day, rarely go to playground, watch TV 2h/day (-0.5 SDs)
  - High: read every day, go to playground once every fortnight, watch TV 1h/day (+0.5 SDs)
- Late Investments (age 10)
  - Low: moderate interest in child's education, don't talk much to child, sometimes engage in activities together (-0.5 SDs)
  - High: a lot of interest in child's education, talk to child a lot, often engage in activities together (+0.5 SDs)

Eliciting parental beliefs

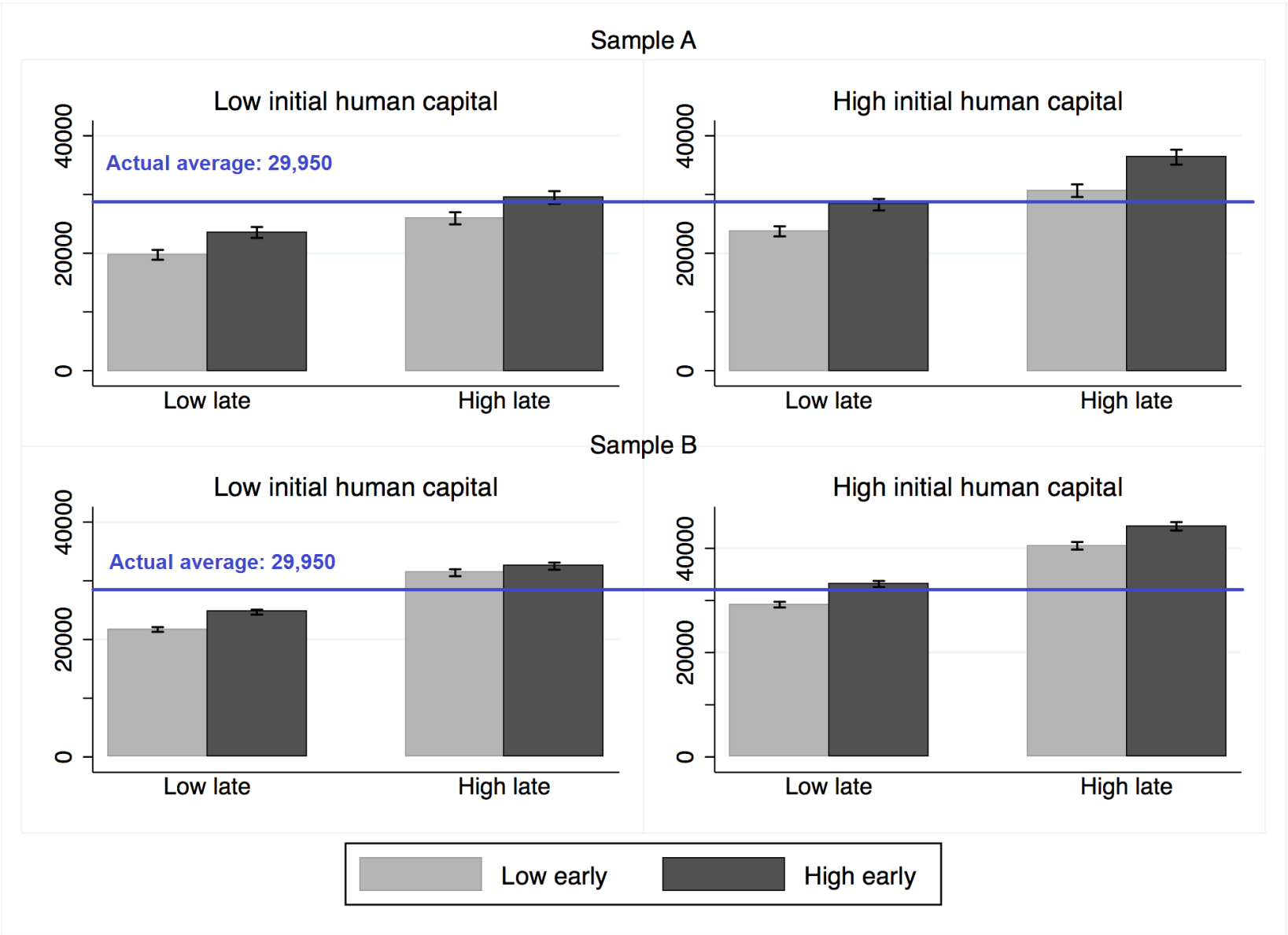
Types and Levels of Inputs (Sample B):

- Initial Human Capital (age 5)
  - Low: scored worse than 70% on intelligence test
  - High: scored better than 70% on intelligence test
- Early Investments (age 5)
  - Low: read every 2nd day, rarely go to playground, watch TV 2h/day (-0.5 SDs)
  - High: read every day, go to playground once every fortnight, watch TV 1h/day (+0.5 SDs)
- Late Investments (age 10)
  - Low: moderate interest in child's education, don't talk much to child, sometimes engage in activities together (-0.5 SDs)
  - High: a lot of interest in child's education, talk to child a lot, often engage in activities together (+0.5 SDs)

Results - Beliefs about production technology



Results - Beliefs about production technology



Results - Beliefs about production technology

Dependent variable: Perceived log earnings at age 30

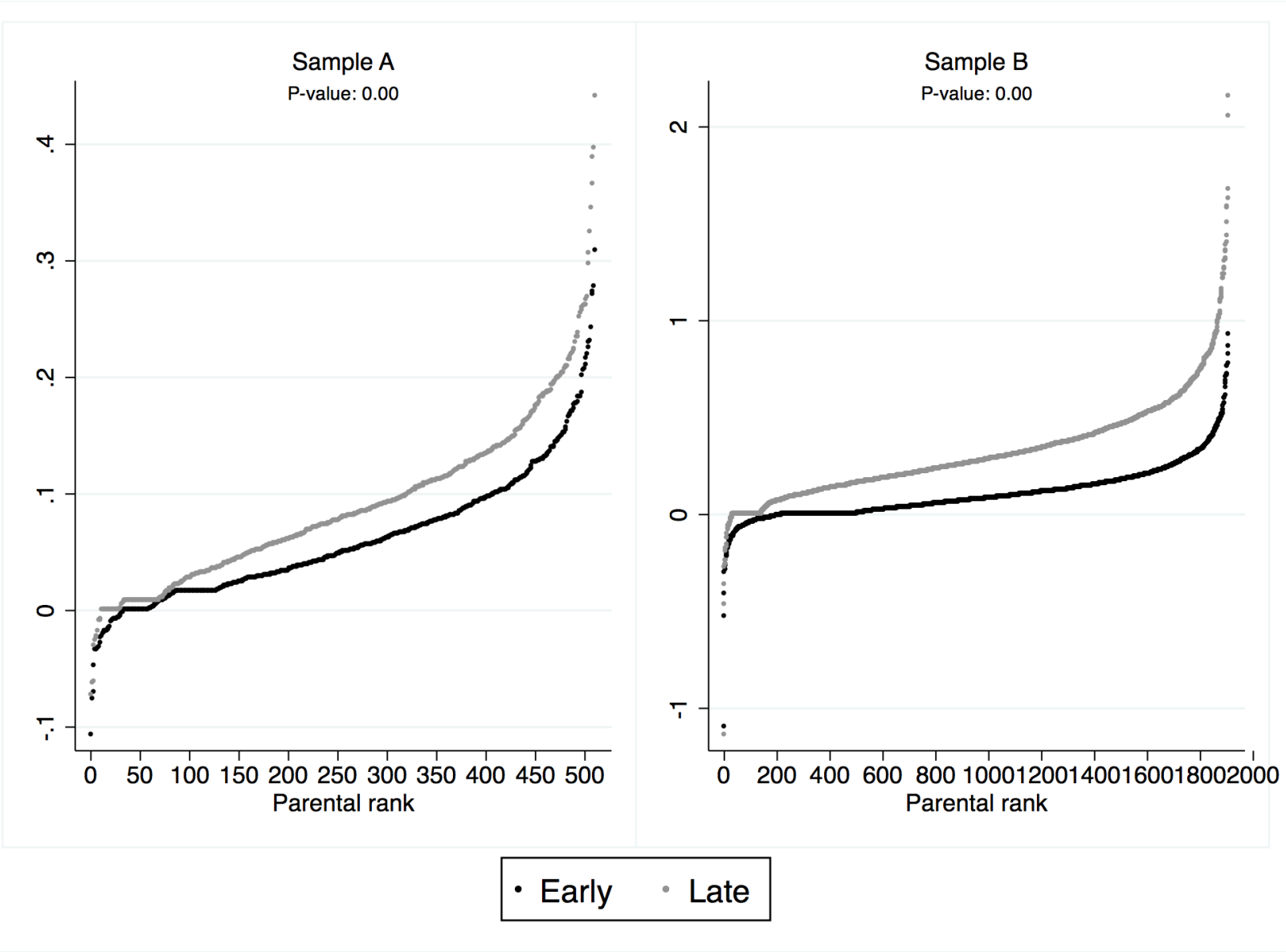
	Sample A		Sample B	
	(1)	(2)	(3)	(4)
Early	0.051***		0.100***	
	(0.003)		(0.003)	
Late	0.085***		0.315***	
	(0.003)		(0.006)	
High HC	0.185***		0.288***	
	(0.010)		(0.006)	
Early x Late				
Early x High HC				
Late x High HC				
Parent fixed effects	Yes		Yes	
Observations	4069		16251	
R^2	0.827		0.782	

Results - Beliefs about production technology

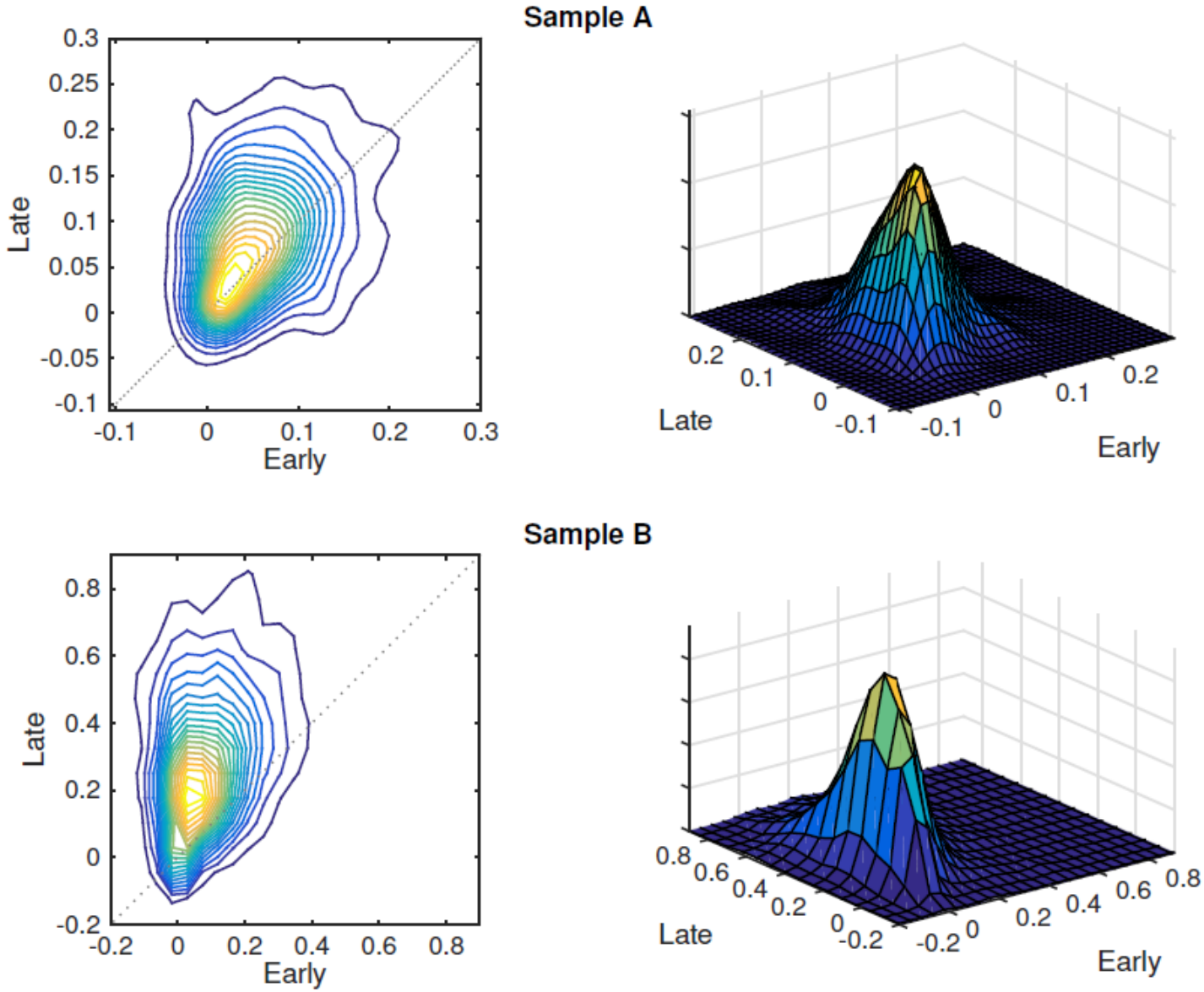
Dependent variable: Perceived log earnings at age 30

	Sample A		Sample B	
	(1)	(2)	(3)	(4)
Early	0.051***	0.060***	0.100***	0.129***
	(0.003)	(0.004)	(0.003)	(0.005)
Late	0.085***	0.096***	0.315***	0.364***
	(0.003)	(0.005)	(0.006)	(0.009)
High HC	0.185***	0.187***	0.288***	0.288***
	(0.010)	(0.016)	(0.006)	(0.008)
Early x Late		-0.005***		-0.077***
		(0.001)		(0.006)
Early x High HC		0.002		0.018***
		(0.003)		(0.005)
Late x High HC		-0.003		-0.019***
		(0.004)		(0.006)
Parent fixed effects	Yes	Yes	Yes	Yes
Observations	4069	4069	16251	16251
R^2	0.827	0.827	0.782	0.784

Results - Heterogeneity



Results - Perceived returns on LHS



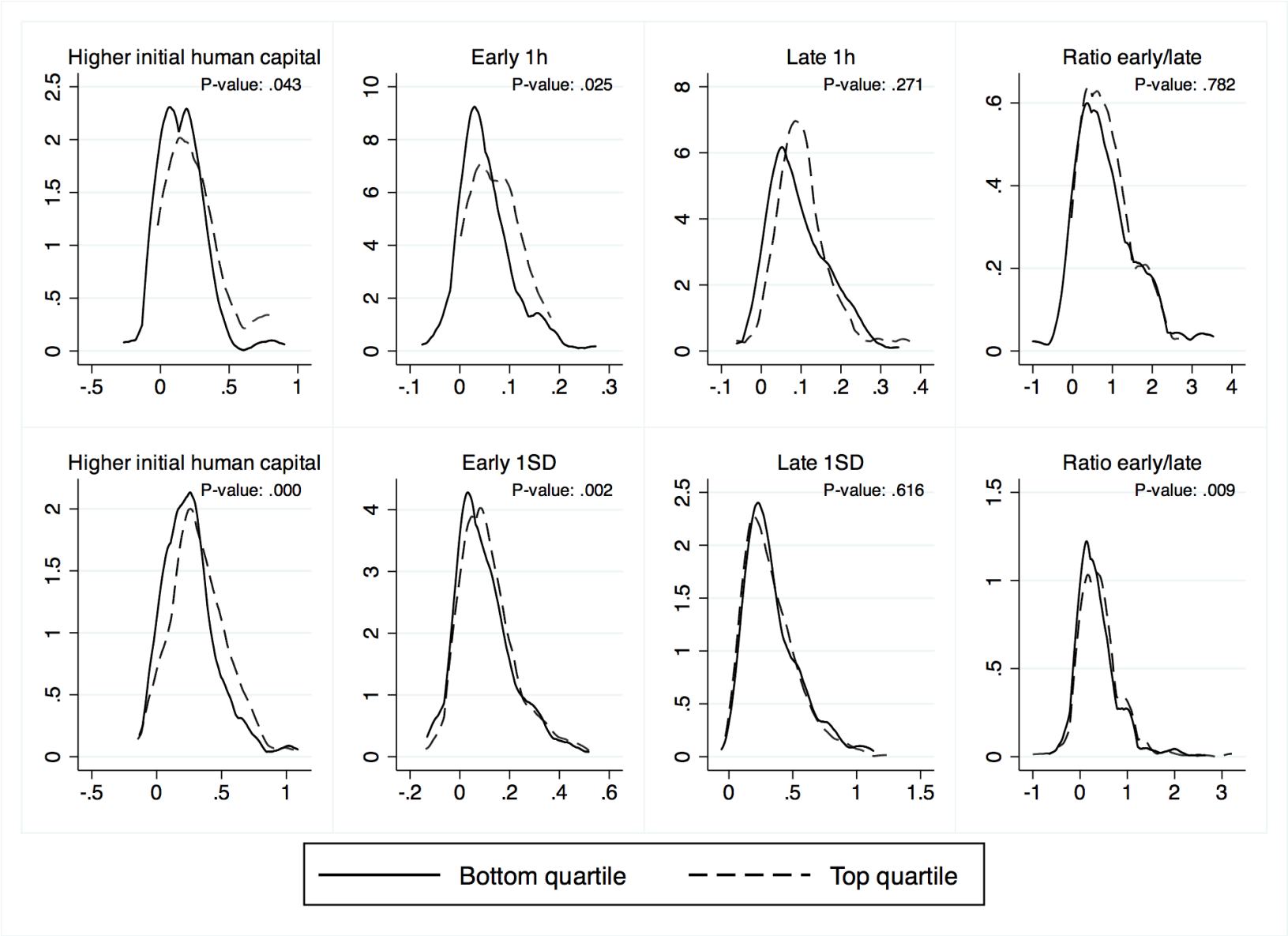
Results - Perceived returns on LHS

	Sample A			
	HC	Early	Late	Ratio
2nd income quartile	0.034	0.006	0.002	0.131
	(0.024)	(0.007)	(0.009)	(0.211)
3rd income quartile	0.020	0.020**	0.020*	-0.025
	(0.029)	(0.009)	(0.011)	(0.249)
4th income quartile	0.073**	0.019*	0.019	0.346
	(0.034)	(0.010)	(0.013)	(0.295)
University graduate	0.007	0.005	-0.009	0.052
	(0.018)	(0.005)	(0.007)	(0.156)
Observations	470	474	474	449
Controls	YES	YES	YES	YES
R^2	0.044	0.025	0.021	0.012

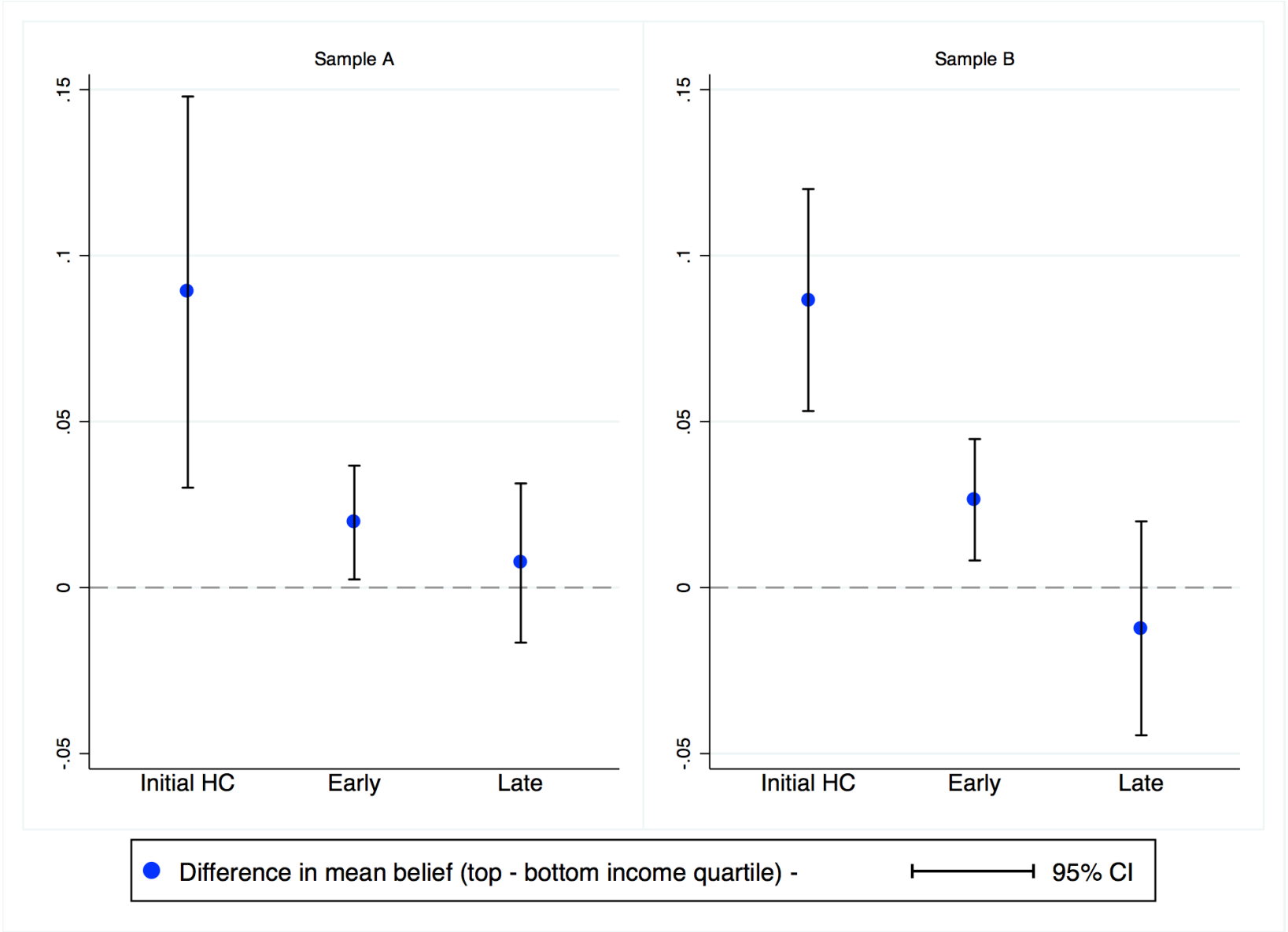
Results - Perceived returns on LHS

	Sample B			
	HC	Early	Late	Ratio
2nd income quartile	0.024	-0.012	-0.013	-0.058*
	(0.018)	(0.009)	(0.017)	(0.034)
3rd income quartile	0.050**	0.018*	0.012	0.039
	(0.020)	(0.010)	(0.019)	(0.036)
4th income quartile	0.073***	0.024**	0.001	0.038
	(0.019)	(0.010)	(0.018)	(0.035)
University graduate	0.028**	0.006	-0.016	0.046*
	(0.014)	(0.007)	(0.013)	(0.026)
Observations	1682	1682	1682	1553
Controls	YES	YES	YES	YES
R^2	0.030	0.017	0.007	0.023

Results - Kernel densities by income quartile



Results - Difference by income quartile



Results - Kernel densities by randomization