Subjective Beliefs and Investment in Early Childhood

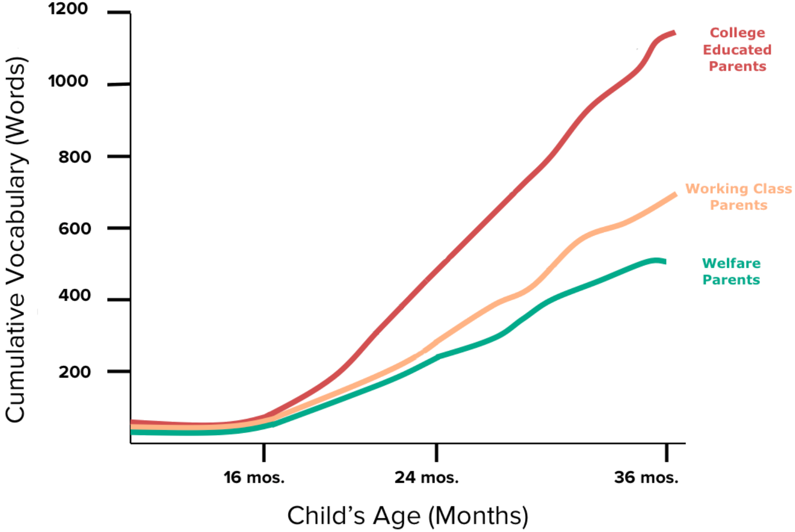
Marcos Lee

Queen's University

EEA-ESEM 2024

August 29, 2024

Investment gaps are substantial across families



Source: Hart and Risley (1995)

What explains investment gaps?

- Credit constraints, time constraints, parental education, . . . (Carneiro and Ginja (2016); Caucutt et al. (2020); Dahl and Lochner (2012); Currie and Moretti (2003); Aizer and Stroud (2010))
  - These factors have poor explanatory power for the observed gaps in speech

What explains investment gaps?

- Credit constraints, time constraints, parental education, . . . (Carneiro and Ginja (2016); Caucutt et al. (2020); Dahl and Lochner (2012); Currie and Moretti (2003); Aizer and Stroud (2010))
    - These factors have poor explanatory power for the observed gaps in speech
- More recently, parental (mis)-beliefs about the returns to investment

What explains investment gaps?

- Credit constraints, time constraints, parental education, . . . (Carneiro and Ginja (2016); Caucutt et al. (2020); Dahl and Lochner (2012); Currie and Moretti (2003); Aizer and Stroud (2010))
  - These factors have poor explanatory power for the observed gaps in speech
- More recently, parental (mis)-beliefs about the returns to investment
- Low income families underestimate returns to investment (Cunha et al. (2013,2020); Boneva and Rauh (2018); Attanasio et al. (2019))

Mean beliefs are an important mechanism...

- Heterogeneity of **mean beliefs** can explain why some families invest more than others (Cunha et al., 2022; Attanasio et al., 2019)

Mean beliefs are an important mechanism...

- Heterogeneity of **mean beliefs** can explain why some families invest more than others (Cunha et al., 2022; Attanasio et al., 2019)
- Increasingly more interventions that educate parents on the importance of investment
  - Jamaica Home Visiting Program; Nurse-Family Partnership Program; Growth Mindset (Grantham-McGregor et al., 1991; Heckman et al., 2017; Rowe and Leech, 2019)
  - Many RCTs on home based interventions (Baranov et al., 2020; Attanasio et al., 2020; List et al., 2021)

But what about belief uncertainty?

- Lot of ground work on **mean beliefs**, but no attention to **belief uncertainty**

But what about belief uncertainty?

- Lot of ground work on **mean beliefs**, but no attention to **belief uncertainty**
- It is an important factor in decision making under uncertainty in a variety of educational contexts

But what about belief uncertainty?

- Lot of ground work on **mean beliefs**, but no attention to **belief uncertainty**
- It is an important factor in decision making under uncertainty in a variety of educational contexts
- It is not measured as there is no methodology to do so (in early childhood contexts)

But what about belief uncertainty?

- Lot of ground work on **mean beliefs**, but no attention to **belief uncertainty**
- It is an important factor in decision making under uncertainty in a variety of educational contexts
- It is not measured as there is no methodology to do so (in early childhood contexts)
- Could be an important driver of why some information interventions work while others don't

This paper

1. Develop a methodology to elicit both mean beliefs and belief uncertainty about the returns to investment in early childhood
   - Estimate belief distributions for each individual parent
   - Measurement error is an important factor that should be dealt with

This paper

1. Develop a methodology to elicit both mean beliefs and belief uncertainty about the returns to investment in early childhood
    - Estimate belief distributions for each individual parent
    - Measurement error is an important factor that should be dealt with
2. Investigate the relationship between beliefs and investment
    - Estimate a model of parental investment that uses the full distribution of beliefs
    - Does belief uncertainty affect investment?

Literature Review

- **Subjective Expectations:** Manski (1993, 2004); Dominitz and Manski (1996); Delavande (2008); Jensen (2010); Stinebrickner and Stinebrickner (2012); Zafar (2011); Delavande and Zafar (2018); Wiswall and Zafar (2015)
- **Parental Information and Investment in Children:** Cunha et al. (2013, 2022); Giustinelli (2016); Boneva and Rauh (2018); Attanasio et al. (2019); Dizon-Ross (2019); List et al. (2021)
- **Uncertainty in Parental Investment:** Attanasio and Kaufmann (2014); Carneiro and Ginja (2016); Tabetando (2019); Sovero (2018); Tanaka and Yamano (2015); Basu and Dimova (2022); Conti et al. (2022), Abbott (2022)
- **Time Investment and Costs:** Del Boca, Flinn, Wiswall (2014,2016); Guryan et al. (2008); Folbre et al. (2005); Kalil et al. (2012); Price (2008); Bono et al. (2016); ; Schoonbroodt (2018)

Outline

1. How to identify and elicit subjective beliefs about the returns to investment in early childhood

2. Survey Instruments

3. Data and Results

Example - Setting and Assumptions

Assume the structural skill production function is given by:

$$\theta_{1,i} = \delta_0 + \delta_1 x_i + \xi_i$$

Example - Setting and Assumptions

Assume the structural skill production function is given by:

$$\theta_{1,i} = \delta_0 + \delta_1 x_i + \xi_i$$

Parents know:

- Their own investment x_i
- The functional form

Example - Setting and Assumptions

Assume the structural skill production function is given by:

$$\theta_{1,i} = \delta_0 + \delta_1 x_i + \xi_i$$

Parents know:

- Their own investment x_i

- The functional form

Parents don't know: δ_0, δ_1, ξ_i

Example - Setting and Assumptions

Assume the structural skill production function is given by:

$$\theta_{1,i} = \delta_0 + \delta_1 x_i + \xi_i$$

Parents know:

- Their own investment x_i

- The functional form

Parents don't know: δ_0, δ_1, $\xi_i \rightarrow \theta_{1,i}$

Example - Setting and Assumptions

Assume the structural skill production function is given by:

$$\theta_{1,i} = \delta_0 + \delta_1 x_i + \xi_i$$

Parents know:

- Their own investment x_i

- The functional form

Parents don't know: δ_0, δ_1, $\xi_i \rightarrow \theta_{1,i}$

- δ_k, $\theta_{1,i}$ are random variables with distribution $G_i^k(\cdot)$

Example - Setting and Assumptions

Assume the structural skill production function is given by:

$$\theta_{1,i} = \delta_0 + \delta_1 x_i + \xi_i$$

Parents know:

- Their own investment x_i
- The functional form

Parents don't know: δ_0, δ_1, $\xi_i \rightarrow \theta_{1,i}$

- δ_k, $\theta_{1,i}$ are random variables with distribution $G_i^k(\cdot)$
- μ_{i,δ_k} and σ^2_{i,δ_k} are the expected belief and uncertainty about δ_k

Example - Belief's Structural Function

Ω_i: Information set of parent i

ξ_i: $E[\xi_i|\Omega_i] = 0$ and $\xi_i \perp \delta_k$

$$\begin{aligned} E[\theta_{i,1}|\Omega_i] &= E[\delta_0 + \delta_1 x_i + \xi_i|\Omega_i] \\ &= \mu_{i,\delta_0} + \mu_{i,\delta_1} x_i, \end{aligned}$$

$$\begin{aligned} Var(\theta_{i,1}|\Omega_i) &= Var(\delta_0 + \delta_1 x_i + \xi_i|\Omega_i) \\ &= (\sigma^2_{i,\delta_0} + \sigma^2_{i,\xi}) + \sigma^2_{i,\delta_1}(x_i)^2 \end{aligned}$$

Example - How to elicit beliefs from parents?

$$\mu_{\theta_{i,1}} \equiv E[\theta_{i,1}|\Omega_i] = \mu_{i,\delta_0} + \mu_{i,\delta_1} x_i,$$
$$\sigma^2_{\theta_{i,1}} \equiv Var(\theta_{i,1}|\Omega_i) = \sigma^2_{i,0} + \sigma^2_{i,\delta_1} x_i^2.$$

- Goal: Identify $\mu_{\delta_k}, \sigma^2_{\delta_k}$

Example - How to elicit beliefs from parents?

$$\mu_{\theta_{i,1}} \equiv E[\theta_{i,1}|\Omega_i] = \mu_{i,\delta_0} + \mu_{i,\delta_1} x_i,$$
$$\sigma^2_{\theta_{i,1}} \equiv Var(\theta_{i,1}|\Omega_i) = \sigma^2_{i,0} + \sigma^2_{i,\delta_1} x_i^2.$$

- Goal: Identify $\mu_{\delta_k}, \sigma^2_{\delta_k}$
- Data: $Z_i = \{x_i\}$ and $(\mu_{\theta_{i,1}}, \sigma^2_{\theta_{i,1}})$

Example - How to elicit beliefs from parents?

$$\mu_{\theta_{i,1}} \equiv E[\theta_{i,1}|\Omega_i] = \mu_{i,\delta_0} + \mu_{i,\delta_1} x_i,$$
$$\sigma^2_{\theta_{i,1}} \equiv Var(\theta_{i,1}|\Omega_i) = \sigma^2_{i,0} + \sigma^2_{i,\delta_1} x_i^2.$$

- Goal: Identify $\mu_{\delta_k}, \sigma^2_{\delta_k}$
- Data: $Z_i = \{x_i\}$ and $(\mu_{\theta_{i,1}}, \sigma^2_{\theta_{i,1}})$
- Key idea: Exogenously vary x_i and collect $\mu_{\theta_{i,1}}$ and $\sigma^2_{\theta_{i,1}}$

Example - How to elicit beliefs from parents?

$$\mu_{\theta_{i,1}} \equiv E[\theta_{i,1}|\Omega_i] = \mu_{i,\delta_0} + \mu_{i,\delta_1} x_i,$$
$$\sigma^2_{\theta_{i,1}} \equiv Var(\theta_{i,1}|\Omega_i) = \sigma^2_{i,0} + \sigma^2_{i,\delta_1} x_i^2.$$

- Goal: Identify $\mu_{\delta_k}, \sigma^2_{\delta_k}$
- Data: $Z_i = \{x_i\}$ and $(\mu_{\theta_{i,1}}, \sigma^2_{\theta_{i,1}})$
- Key idea: Exogenously vary x_i and collect $\mu_{\theta_{i,1}}$ and $\sigma^2_{\theta_{i,1}}$
- Assume for now that we can directly ask individuals to provide $\mu_{\theta_{i,1}}$ and $\sigma^2_{\theta_{i,1}}$

Example - Scenarios

- Establish two hypothetical scenarios of investment x_i: High ($\bar{x}$) and Low ($\underline{x}$)

Example - Scenarios

- Establish two hypothetical scenarios of investment x_i: High ($\bar{x}$) and Low ($\underline{x}$)
- Respondents are asked to provide $\mu_{\theta_{i,1}}$ and $\sigma^2_{\theta_{i,1}}$:
  - For $\bar{x}$: $\bar{\mu}_{\theta_{i,1}}$ and $\bar{\sigma}^2_{\theta_{i,1}}$
  - For $\underline{x}$: $\underline{\mu}_{\theta_{i,1}}$ and $\underline{\sigma}^2_{\theta_{i,1}}$

Example - Identification from moments

$$\bar{\mu}_{\theta_{i,1}} - \underline{\mu}_{\theta_{i,1}} = E[\theta_{i,1}|\Omega_i, \bar{x}] - E[\theta_{i,1}|\Omega_i, \underline{x}]$$
$$= \mu_{i,\delta_1}(\bar{x} - \underline{x}),$$

$$\bar{\sigma}^2_{\theta_{i,1}} - \underline{\sigma}^2_{\theta_{i,1}} = Var(\theta_{i,1}|\Omega_i, \bar{x}) - Var(\theta_{i,1}|\Omega_i, \underline{x})$$
$$= \sigma^2_{i,\delta_1}(\bar{x}^2 - \underline{x}^2),$$

Example - Identification and Measurement Error

- Parents provide $\mu_{\theta_{i,1}}$ and $\sigma^2_{\theta_{i,1}}$ for **different scenarios** (J)

Example - Identification and Measurement Error

- Parents provide $\mu_{\theta_{i,1}}$ and $\sigma^2_{\theta_{i,1}}$ for **different scenarios** (J)

- Overidentified system based on moments $\bar{\mu}_{\theta_i} - \underline{\mu}_{\theta_i}$, $\bar{\sigma}^2_{\theta_i} - \underline{\sigma}^2_{\theta_i}$

Example - Identification and Measurement Error

- Parents provide $\mu_{\theta_{i,1}}$ and $\sigma^2_{\theta_{i,1}}$ for **different scenarios** (J)
- Overidentified system based on moments $\bar{\mu}_{\theta_i} - \underline{\mu}_{\theta_i}$, $\bar{\sigma}^2_{\theta_i} - \underline{\sigma}^2_{\theta_i}$
- However, using these moments assumes there is no measurement error: $\mu_{\theta_{i,1}}$ is a perfect measure of $E[\theta_1|\Omega_i]$

Example - Identification and Measurement Error

- Parents provide $\mu_{\theta_{i,1}}$ and $\sigma^2_{\theta_{i,1}}$ for **different scenarios** (J)
- Overidentified system based on moments $\bar{\mu}_{\theta_i} - \underline{\mu}_{\theta_i}$, $\bar{\sigma}^2_{\theta_i} - \underline{\sigma}^2_{\theta_i}$
- However, using these moments assumes there is no measurement error: $\mu_{\theta_{i,1}}$ is a perfect measure of $E[\theta_1|\Omega_i]$
- $\mu_{\theta_{i,j,1}}$ and $\sigma^2_{\theta_{i,j,1}}$ are **error-ridden** measures of $E[\theta_1|\Omega_i]$ and $Var(\theta_1|\Omega_i)$

Example - Measurement Error Model

$$
\begin{aligned}
\mu_{\theta_{i,j,1}} &= \overbrace{\mu_{i,\delta_0} + \mu_{i,\delta_1} x_j}^{E[\theta_{i,1}|\Omega_i]} \quad + \epsilon_{i,j,1}, \\
\sigma^2_{\theta_{i,j,1}} &= \underbrace{\sigma^2_{i,0} + \sigma^2_{i,\delta_1} x_j^2}_{Var(\theta_{i,1}|\Omega_i)} \quad + \epsilon_{i,j,2}.
\end{aligned}
$$

There are $i = 1, \ldots, N$ individuals, and each individual has J observations.

This is a typical case of a ‘repeated measurements’ model (Schennach, 2016).

System of Random Coefficients Regression Model (Swamy, 1970): efficient individual-level estimates of μ_{δ_k} and $\sigma^2_{\delta_k}$.

Example - How to ask about latent variables?

- From national data (e.g., NLSY), Motor and Social Development (MSD) scale asks, at different ages *a*:
  - Does your child **know how to speak a sentence of 3 or more words?** *YES/NO*

Example - How to ask about latent variables?

- From national data (e.g., NLSY), Motor and Social Development (MSD) scale asks, at different ages *a*:
  - Does your child **know how to speak a sentence of 3 or more words?** *YES/NO*
- Estimate Item Response Theory (IRT) model and obtain the population distribution of skills θ
  - Provides a mapping between a $\hat{\theta}$ and the ages a child speaking a sentence of 3 words
  - If a child speaks a sentence of 3 words at **15 months** $\rightarrow$ High $\hat{\theta}$
  - If a child speaks a sentence of 3 words at **25 months** $\rightarrow$ Low $\hat{\theta}$

Example - Translating parents beliefs abouts ages to skills

Example - Translating parents beliefs abouts ages to skills

- "Flip" MSD scale to ask about beliefs:

Example - Translating parents beliefs abouts ages to skills

- "Flip" MSD scale to ask about beliefs:
  - Imagine a parent that has a *HIGH* investment in their child:
  - What do you think are the *YOUNGEST*, *MOST LIKELY*, *OLDEST* ages the child will **learn to speak a sentence of 3 or more words**?

Example - Translating parents beliefs abouts ages to skills

- "Flip" MSD scale to ask about beliefs:
  - Imagine a parent that has a *HIGH* investment in their child:
  - What do you think are the *YOUNGEST*, *MOST LIKELY*, *OLDEST* ages the child will **learn to speak a sentence of 3 or more words**?
- From $\underline{a}$, $\dot{a}$, $\bar{a}$: Convert these values using IRT
- Have a measure of beliefs that is anchored on a nationally representative distribution of skills

Actual Setting and Model

- Cobb-Douglas production function:

$$\ln \theta_{1,i} = \delta_0 + \delta_1 \ln \theta_{0,i} + \delta_2 \ln x_i + \xi_i$$

Actual Setting and Model

- Cobb-Douglas production function:

$$\ln \theta_{1,i} = \delta_0 + \delta_1 \ln \theta_{0,i} + \delta_2 \ln x_i + \xi_i$$

- 4 activities L and 4 scenarios J for each i: 16 repeated measures
  - Speak a sentence of 3 words; Counts 3 Objects; Says First and Last Name; Know age and sex (MSD Instrument)
  - High/Low investment (x_i); Normal/Poor health at birth ($\theta_{0,i}$)

Actual Setting and Model

- Cobb-Douglas production function:

$$\ln \theta_{1,i} = \delta_0 + \delta_1 \ln \theta_{0,i} + \delta_2 \ln x_i + \xi_i$$

- 4 activities L and 4 scenarios J for each i: 16 repeated measures
  - Speak a sentence of 3 words; Counts 3 Objects; Says First and Last Name; Know age and sex (MSD Instrument)
  - High/Low investment (x_i); Normal/Poor health at birth ($\theta_{0,i}$)
- 7 parameters to be identified for each i: μ_{i,δ_k} and σ^2_{i,δ_k}

Survey Instrument - Scenarios

1) A **"normal health"** baby is one whose gestation lasted 9 months, weighed 8 pounds, and measured 20 inches at birth. A **"poor health"** baby is one whose gestation lasted 7 months, weighed 5 pounds, and measured 18 inches at birth.

2) A **"high intensity"** interaction is one in which the mothers spends 6 hours a day with the baby in active interaction, while a **"low intensity"** one the mother spends 2 hours a day with the baby in active interaction. These interactions includes activities such as:

(a) soothing the baby when he/she is upset;
(b) moving the baby's arms and legs around playfully;
(c) playing peek-a-boo with the baby;
(d) singing songs with the baby;
(e) speaking to the baby;
(f) feeding, nursing, bathing, attending to health needs;
(...)

Survey Instrument

Please consider a baby with **"normal health"** and a **"high intensity"** interaction between mother and baby.

What do you think is the youngest, most likely, and oldest age a baby learns to **speak a partial sentence of 3 or more words (for example, "Mommy get in car", "Me go too", "No more juice", "All done now")**?

Age in months

0 2 5 7 10 12 14 17 19 22 24 26 29 31 34 36 38 41 43 46 48

Youngest Age (Months)

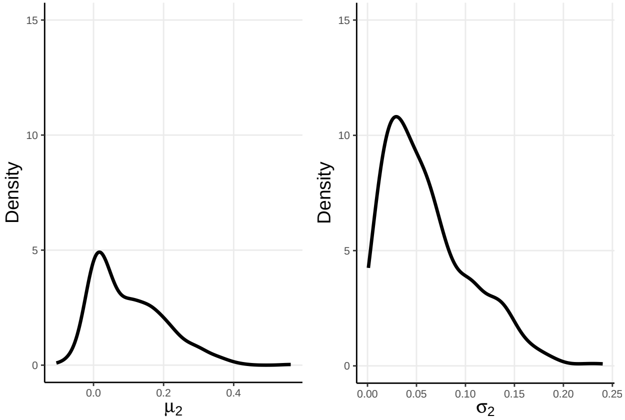
Most Likely Age (Months)

Oldest Age (Months)

Data

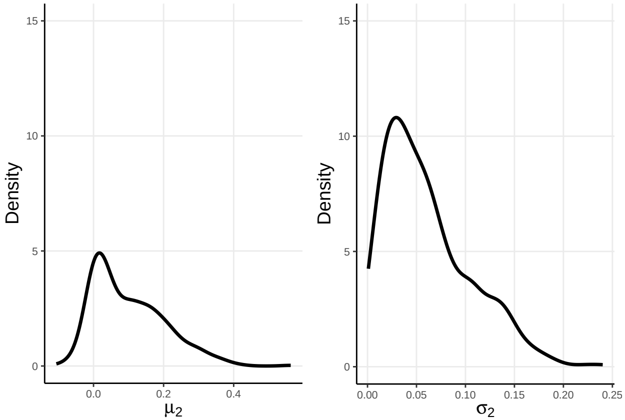
- Qualtrics data collection: 732 women, ages 18-40, with at least one child, oldest under 5 years old
- Collect subjective beliefs, socio-economic variables, and investment measures in own child
- Sample slighly over-represented by Hispanic and Black, lower income
- Internal consistency tests + Correlations consistent with literature

Estimates of μ_{i,δ_2} and σ_{i,δ_2}



- $\hat{\mu}_{\delta_2}$: 0.101 - return is quite low; objective returns estimated to be between 0.2-0.3

Estimates of μ_{i,δ_2} and σ_{i,δ_2}



- $\hat{\sigma}^2_{\delta_2}$: 0.05 - Low uncertainty

How are beliefs correlated with demographics?

- Non-Hispanic Black have **lower mean** and **more uncertain**

How are beliefs correlated with demographics?

- Non-Hispanic Black have **lower mean** and **more uncertain**

- Younger, single, and with more children are **more uncertain**

How are beliefs correlated with demographics?

- Non-Hispanic Black have **lower mean** and **more uncertain**
- Younger, single, and with more children are **more uncertain**
- Higher education $->$ **higher mean**

An Application to a Model With Reference Dependent Preferences

- Model of investment with reference dependent preferences
  - Piloting of survey shows that parents use developmental benchmarks as reference points
  - In general not worried about variances

$$\begin{aligned} u_i(c_i, h_{l_i}, \theta_{i,1}) =& \alpha_1 \ln c_i + \alpha_2 \ln h_{l_i} + \alpha_3 \ln \theta_{i,1} + \\ & \alpha_4 (\ln \theta_{i,1} - \ln \theta_{ref}) \mathbb{1}\{(\ln \theta_{i,1} \leq \ln \theta_{ref})\} \end{aligned}$$

An Application to a Model With Reference Dependent Preferences

- Model of investment with reference dependent preferences
  - Piloting of survey shows that parents use developmental benchmarks as reference points
  - In general not worried about variances

$$
\begin{aligned}
u_i(c_i, h_{l_i}, \theta_{i,1}) = & \alpha_1 \ln c_i + \alpha_2 \ln h_{l_i} + \alpha_3 \ln \theta_{i,1} + \\
& \alpha_4 (\ln \theta_{i,1} - \ln \theta_{ref}) \mathbb{1}\{(\ln \theta_{i,1} \leq \ln \theta_{ref})\}
\end{aligned}
$$

- My measure of skill is the developmental age of the child at 24 months
- Natural reference points are developmental age benchmarks, e.g. $\theta_{ref} = 18$

Impact of Changing Beliefs on Investment

	$\theta_{ref} = 18$	$\theta_{ref} = 21$
10% μ_{i,δ_2}	1.80%	2.70%
10% σ_{i,δ_2}	2.58%	1.79%

- Increasing mean beliefs increases investment

Impact of Changing Beliefs on Investment

	$\theta_{ref} = 18$	$\theta_{ref} = 21$
10% μ_{i,δ_2}	1.80%	2.70%
10% σ_{i,δ_2}	2.58%	1.79%

- Increasing mean beliefs increases investment
- Increasing uncertainty also increases investment!

Impact of Changing Beliefs on Investment

	$\theta_{ref} = 18$	$\theta_{ref} = 21$
10% μ_{i,δ_2}	1.80%	2.70%
10% σ_{i,δ_2}	2.58%	1.79%

- Increasing mean beliefs increases investment
- Increasing uncertainty also increases investment!
  - Increase in investment is driven by those who hold low mean beliefs
  - Consistent with individuals wanting to move away from the reference point
  - Indeed, increase in investment is lower at higher reference points!

Conclusion

Main Contributions

- I develop a methodology to elicit subjective belief distribution about returns to investment

- I also elicit the subjective price of investment

- I show how we can use these beliefs to estimate a model of investment