Identifying Procrastination from the Timing of Choices

Paul Heidhues Philipp Strack

[1]DICE, University of Düsseldorf

[2]Yale

September 4.

Motivation

Time-inconsistency: intuition and evidence suggests that people are time-inconsistent and present-focused.

- Strotz (1956); Frederick et al. (2002); Augenblick et al. (2015); Augenblick and Rabin (2019).

Motivation

Time-inconsistency: intuition and evidence suggests that people are time-inconsistent and present-focused.

- Strotz (1956); Frederick et al. (2002); Augenblick et al. (2015); Augenblick and Rabin (2019).
- Degree of time-inconsistency important for optimal policy design.

Motivation

Time-inconsistency: intuition and evidence suggests that people are time-inconsistent and present-focused.

- Strotz (1956); Frederick et al. (2002); Augenblick et al. (2015); Augenblick and Rabin (2019).
- Degree of time-inconsistency important for optimal policy design.

Psychological intuition: time-inconsistent preferences—especially in combination with forecasting errors about own future behavior—leads to procrastination (tendency to delay unpleasant tasks).

Motivation

Time-inconsistency: intuition and evidence suggests that people are time-inconsistent and present-focused.

- Strotz (1956); Frederick et al. (2002); Augenblick et al. (2015); Augenblick and Rabin (2019).
- Degree of time-inconsistency important for optimal policy design.

Psychological intuition: time-inconsistent preferences—especially in combination with forecasting errors about own future behavior—leads to procrastination (tendency to delay unpleasant tasks). Formalized, in

- Akerlof (1991), and
- O'Donoghue and Rabin (1999a,b, 2001).

Motivation

Stylized fact: in line with procrastination, people often complete tasks last minute.

Motivation

Stylized fact: in line with procrastination, people often complete tasks last minute.

- Homework, EARIE conference registrations.

Motivation

Stylized fact: in line with procrastination, people often complete tasks last minute.

- Homework, EARIE conference registrations.
- Parking tickets (Heffetz et al., 2016); health care plan choice (Brown and Previtero, 2018); taxes (e.g Martinez et al., 2017); patent officers' fillings (Frakes and Wasserman, 2016).

Motivation

Stylized fact: in line with procrastination, people often complete tasks last minute.

- Homework, EARIE conference registrations.
- Parking tickets (Heffetz et al., 2016); health care plan choice (Brown and Previtero, 2018); taxes (e.g Martinez et al., 2017); patent officers' fillings (Frakes and Wasserman, 2016).

Natural idea: if task completion is driven by the tendency to procrastinate, use data on task completion to identify time preferences.

Motivation

Stylized fact: in line with procrastination, people often complete tasks last minute.

- Homework, EARIE conference registrations.
- Parking tickets (Heffetz et al., 2016); health care plan choice (Brown and Previtero, 2018); taxes (e.g Martinez et al., 2017); patent officers' fillings (Frakes and Wasserman, 2016).

Natural idea: if task completion is driven by the tendency to procrastinate, use data on task completion to identify time preferences.

- "Common wisdom" (see, e.g., Frakes and Wasserman, 2016): observed bunching at the deadline is evidence of time-inconsistency.
- Argument: inconsistent with $\delta \approx 1$.

Motivation

Stylized fact: in line with procrastination, people often complete tasks last minute.

- Homework, EARIE conference registrations.
- Parking tickets (Heffetz et al., 2016); health care plan choice (Brown and Previtero, 2018); taxes (e.g Martinez et al., 2017); patent officers' fillings (Frakes and Wasserman, 2016).

Natural idea: if task completion is driven by the tendency to procrastinate, use data on task completion to identify time preferences.

- "Common wisdom" (see, e.g., Frakes and Wasserman, 2016): observed bunching at the deadline is evidence of time-inconsistency.
- Argument: inconsistent with $\delta \approx 1$.

Option-value of waiting: if task completion cost are not deterministic, may wait for lower cost draw.

- See Wald (1945); Weisbrod (1964); Dixit and Pindyck (1994).
- Waiting decision determined by other factors than time-preferences!

Motivation

Stylized fact: in line with procrastination, people often complete tasks last minute.

- Homework, EARIE conference registrations.
- Parking tickets (Heffetz et al., 2016); health care plan choice (Brown and Previtero, 2018); taxes (e.g Martinez et al., 2017); patent officers' fillings (Frakes and Wasserman, 2016).

Natural idea: if task completion is driven by the tendency to procrastinate, use data on task completion to identify time preferences.

- "Common wisdom" (see, e.g., Frakes and Wasserman, 2016): observed bunching at the deadline is evidence of time-inconsistency.
- Argument: inconsistent with $\delta \approx 1$.

Option-value of waiting: if task completion cost are not deterministic, may wait for lower cost draw.

- See Wald (1945); Weisbrod (1964); Dixit and Pindyck (1994).
- Waiting decision determined by other factors than time-preferences!
- One way to disentangle: parameteric assumptions on net-benefit or opportunity cost distribution (Martinez et al., 2017).

Uncovering Time-Preferences

Our research question: can one **non-parameterically** uncover time preference from task completion data?

Our research question: can one **non-parameterically** uncover time preference from task completion data?

- We suppose the net benefit/opportunity cost of doing the task at a given point in time are unobservable to the analyst.

Our research question: can one **non-parameterically** uncover time preference from task completion data?

- We suppose the net benefit/opportunity cost of doing the task at a given point in time are unobservable to the analyst.
- Otherwise, we impose **a lot** of helpful structure.

Uncovering Time-Preferences

Our research question: can one **non-parameterically** uncover time preference from task completion data?

- We suppose the net benefit/opportunity cost of doing the task at a given point in time are unobservable to the analyst.
- Otherwise, we impose **a lot** of helpful structure.

In a nutshell: Despite strong stationary, homogeneity, and observability assumptions, and restriction to quasi-hyperbolic discounting, we find:

- The degree of present bias and the discount factor are **for any data set** unidentifiable in our basic setting.

Uncovering Time-Preferences

Our research question: can one **non-parameterically** uncover time preference from task completion data?

- We suppose the net benefit/opportunity cost of doing the task at a given point in time are unobservable to the analyst.
- Otherwise, we impose **a lot** of helpful structure.

In a nutshell: Despite strong stationary, homogeneity, and observability assumptions, and restriction to quasi-hyperbolic discounting, we find:

- The degree of present bias and the discount factor are **for any data set** unidentifiable in our basic setting.
- Importantly, present bias parameter is unidentified even when fixing the long-run discount factor.

Our research question: can one **non-parameterically** uncover time preference from task completion data?

- We suppose the net benefit/opportunity cost of doing the task at a given point in time are unobservable to the analyst.
- Otherwise, we impose **a lot** of helpful structure.

In a nutshell: Despite strong stationary, homogeneity, and observability assumptions, and restriction to quasi-hyperbolic discounting, we find:

- The degree of present bias and the discount factor are **for any data set** unidentifiable in our basic setting.
- Importantly, present bias parameter is unidentified even when fixing the long-run discount factor.
- **Naivite vs Sophistication** are also not identifiable.

Uncovering Time-Preferences

Our research question: can one **non-parameterically** uncover time preference from task completion data?

- We suppose the net benefit/opportunity cost of doing the task at a given point in time are unobservable to the analyst.
- Otherwise, we impose **a lot** of helpful structure.

In a nutshell: Despite strong stationary, homogeneity, and observability assumptions, and restriction to quasi-hyperbolic discounting, we find:

- The degree of present bias and the discount factor are **for any data set** unidentifiable in our basic setting.
- Importantly, present bias parameter is unidentified even when fixing the long-run discount factor.
- **Naivite vs Sophistication** are also not identifiable.
- With a stationary net-benefit distribution, a hyperbolic discounter never sets an earlier deadline.

Overview

1. Quasi-Hyperbolic Discounting
2. Task Completion Problem
3. The Analyst's Problem
4. Examples
5. Agent's Behavior
6. Non-Identifiability
7. Non-Identifiability for Naives
8. Non-Identifiability for Sophisticates
9. A Priori Knowledge
10. Rich Data
11. Relationship to Dynamic Discrete Choice Literature

Quasi-Hyperbolic Discounting

Time-Preferenes: Quasi-Hyperbolic Discounting

- Let u_t denote the instantaneous utility in period $t \in \{1, \cdots, T+1\}$.

Time-Preferenes: Quasi-Hyperbolic Discounting

- Let u_t denote the instantaneous utility in period $t \in \{1, \cdots, T+1\}$.
- The agent's time preference satisfy **quasi-hyperbolic discounting**.
  - $\delta \in (0, 1]$ is the long-run discount rate.
  - $\beta \in (0, 1]$ is the present-bias parameter.

Time-Preferenes: Quasi-Hyperbolic Discounting

- Let u_t denote the instantaneous utility in period $t \in \{1, \cdots, T+1\}$.
- The agent's time preference satisfy **quasi-hyperbolic discounting**.
  - $\delta \in (0, 1]$ is the long-run discount rate.
  - $\beta \in (0, 1]$ is the present-bias parameter.
- The agent has a belief $\hat{\beta} \in (0, 1]$ regarding her future selves' taste for immediate gratification (O'Donoghue and Rabin, 1999a).
  - $\hat{\beta} = 1$ is the case of naivete.
  - $\hat{\beta} = \beta$ is the case of sophistication.

Time-Preferenes: Quasi-Hyperbolic Discounting

- Let u_t denote the instantaneous utility in period $t \in \{1, \cdots, T+1\}$.
- The agent's time preference satisfy **quasi-hyperbolic discounting**.
  - $\delta \in (0, 1]$ is the long-run discount rate.
  - $\beta \in (0, 1]$ is the present-bias parameter.
- The agent has a belief $\hat{\beta} \in (0, 1]$ regarding her future selves' taste for immediate gratification (O'Donoghue and Rabin, 1999a).
  - $\hat{\beta} = 1$ is the case of naivete.
  - $\hat{\beta} = \beta$ is the case of sophistication.
- We solve for **perception-perfect equilibria**.

Time-Preferenes: Quasi-Hyperbolic Discounting

- Let u_t denote the instantaneous utility in period $t \in \{1, \cdots, T+1\}$.
- The agent's time preference satisfy **quasi-hyperbolic discounting**.
  - $\delta \in (0,1]$ is the long-run discount rate.
  - $\beta \in (0,1]$ is the present-bias parameter.
- The agent has a belief $\hat{\beta} \in (0,1]$ regarding her future selves' taste for immediate gratification (O'Donoghue and Rabin, 1999a).
  - $\hat{\beta} = 1$ is the case of naivete.
  - $\hat{\beta} = \beta$ is the case of sophistication.
- We solve for **perception-perfect equilibria**. Time t self maximizes

$$U^t = u_t + \beta\, \mathbb{E}\left[\sum_{s=t+1}^{T+1} \delta^{s-t}\ u_s,\right]$$

thinking future selves $r > t$ maximize

$$\hat{U}^r = u_r + \hat{\beta}\, \mathbb{E}\left[\sum_{s=r+1}^{T+1} \delta^{s-r}\ u_s\right].$$

Task Completion Problem

Agent's Task-Completion Problem

- Think of **preparing your taxes** or **paying a parking ticket**.
- The agent needs to complete the task before the deadline T.

Agent's Task-Completion Problem

- Think of **preparing your taxes** or **paying a parking ticket**.
- The agent needs to complete the task before the deadline T.
- If she did not complete the task by the end of period T, the agent gets a penalty of $\underline{y}/(\beta\delta) \leq 0$ in period $T+1$.
  - So $\underline{y}$ is the period-T continuation value when not having done the task.

Agent's Task-Completion Problem

- Think of **preparing your taxes** or **paying a parking ticket**.
- The agent needs to complete the task before the deadline T.
- If she did not complete the task by the end of period T, the agent gets a penalty of $\underline{y}/(\beta\delta) \leq 0$ in period $T+1$.
  - So $\underline{y}$ is the period-T continuation value when not having done the task.
  - Task could be mandatory ($\underline{y} = -\infty$) or voluntary ($\underline{y} = 0$).

Agent's Task-Completion Problem

- Think of **preparing your taxes** or **paying a parking ticket**.
- The agent needs to complete the task before the deadline T.
- If she did not complete the task by the end of period T, the agent gets a penalty of $\underline{y}/(\beta\delta) \leq 0$ in period $T+1$.
  - So $\underline{y}$ is the period-T continuation value when not having done the task.
  - Task could be mandatory ($\underline{y} = -\infty$) or voluntary ($\underline{y} = 0$).
- In every period $t \leq T$, the instantaneous utility of completing the task is drawn independently from a given payoff distribution F.
  - Instantaneous benefit of completing the task net of opportunity costs.
  - F is known to the agent.
  - Instantaneous utility of not doing the task is normalized to zero.

Agent's Task-Completion Problem

- Think of **preparing your taxes** or **paying a parking ticket**.
- The agent needs to complete the task before the deadline T.
- If she did not complete the task by the end of period T, the agent gets a penalty of $\underline{y}/(\beta\delta) \leq 0$ in period $T+1$.
    - So $\underline{y}$ is the period-T continuation value when not having done the task.
    - Task could be mandatory ($\underline{y} = -\infty$) or voluntary ($\underline{y} = 0$).
- In every period $t \leq T$, the instantaneous utility of completing the task is drawn independently from a given payoff distribution F.
    - Instantaneous benefit of completing the task net of opportunity costs.
    - F is known to the agent.
    - Instantaneous utility of not doing the task is normalized to zero.
- Tie-breaking assumption: when indifferent, agent waits.
    - Absent a tie-breaking assumption, can rationalize any data by assuming cost are always zero.

Agent's Task-Completion Problem

- Think of **preparing your taxes** or **paying a parking ticket**.
- The agent needs to complete the task before the deadline T.
- If she did not complete the task by the end of period T, the agent gets a penalty of $\underline{y}/(\beta\delta) \leq 0$ in period $T+1$.
  - So $\underline{y}$ is the period-T continuation value when not having done the task.
  - Task could be mandatory ($\underline{y} = -\infty$) or voluntary ($\underline{y} = 0$).
- In every period $t \leq T$, the instantaneous utility of completing the task is drawn independently from a given payoff distribution F.
  - Instantaneous benefit of completing the task net of opportunity costs.
  - F is known to the agent.
  - Instantaneous utility of not doing the task is normalized to zero.
- Tie-breaking assumption: when indifferent, agent waits.
  - Absent a tie-breaking assumption, can rationalize any data by assuming cost are always zero.
  - Exact tie-breaking rule not important.

The Analyst's Problem

Analyst's Problem.

- The analyst observes agent's stopping probabilities at every point in time.
  - Either observes infinitely many homogeneous agents,
  - or the same agent infinitely many times.

Analyst's Problem.

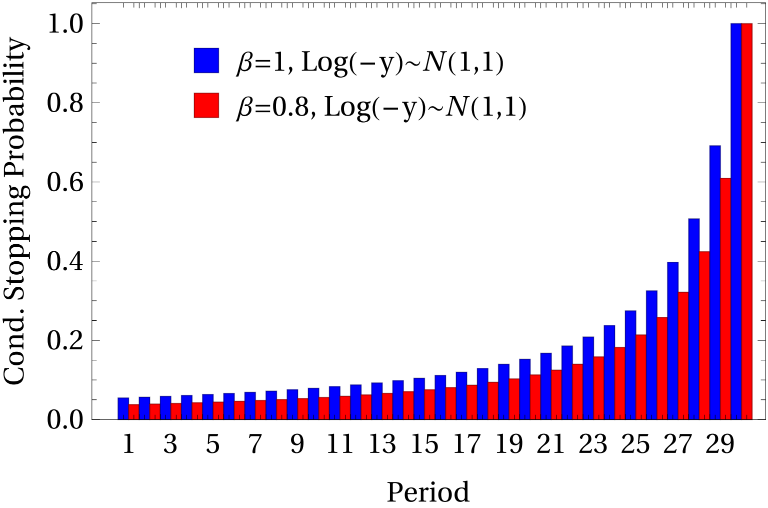
- The analyst observes agent's stopping probabilities at every point in time.
    - Either observes infinitely many homogeneous agents,
    - or the same agent infinitely many times.
    - Obviously homogeneity facilitates identifying time preferences.

Analyst's Problem.

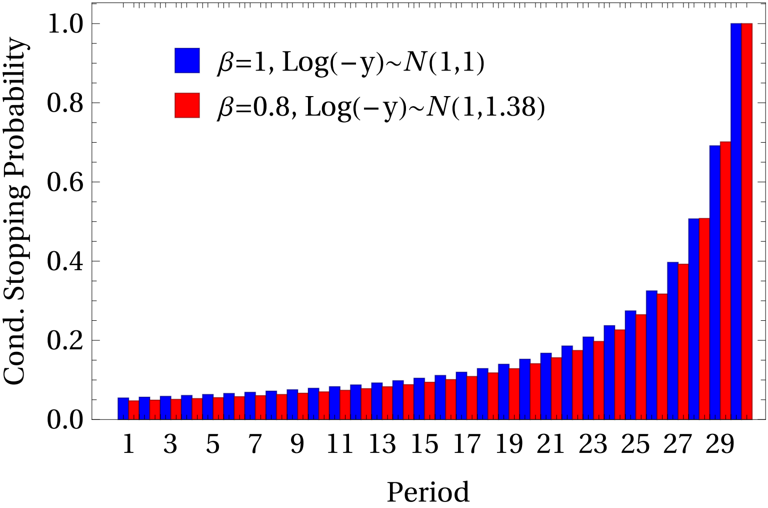
- The analyst observes agent's stopping probabilities at every point in time.
  - Either observes infinitely many homogeneous agents,
  - or the same agent infinitely many times.
  - Obviously homogeneity facilitates identifying time preferences.
- We suppose it is known that opportunity costs are i.i.d.
  - Otherwise can rationalize any data by assuming cost are either one or zero, with the probability that they are zero being equal to a period's stopping probability.
  - Well known in dynamic discrete choice literature (e.g., Section 3.5 in Rust, 1994; Magnac and Thesmar, 2002).
- Best case scenario for identification!

Examples

Observed Task Completion Time for a Mandatory Task



Observed Task Completion Time for a Mandatory Task



Parameter Estimates of β

Agent is time-consistent with $\beta = \delta = 1$, $T = 5$, and the value of completing the task is drawn from a uniform distribution with mean 0 and variance 1.

Parameter Estimates of β

Agent is time-consistent with $\beta = \delta = 1$, $T = 5$, and the value of completing the task is drawn from a uniform distribution with mean 0 and variance 1.

Analyst knows

- the true mean and standard deviation of F
- correctly imposes $\delta = 1$
- but does not know the functional form of F.

Parameter Estimates of β

Agent is time-consistent with $\beta = \delta = 1$, $T = 5$, and the value of completing the task is drawn from a uniform distribution with mean 0 and variance 1.

Analyst knows

- the true mean and standard deviation of F
- correctly imposes $\delta = 1$
- but does not know the functional form of F.

Parameter Estimates of β

Agent is time-consistent with $\beta = \delta = 1$, $T = 5$, and the value of completing the task is drawn from a uniform distribution with mean 0 and variance 1.

Analyst knows

- the true mean and standard deviation of F
- correctly imposes $\delta = 1$
- but does not know the functional form of F.

Parametric Family	*Sq. Distance Minimzation*		*Likelihood Maximization*	
	β	Distance	β	Log-Likelihood
Normal Naive	0.82	0.00231668	0.82	-1.59187
Normal Sophisticate	0.82	0.00267663	0.82	-1.59188
Extreme Value Naive	0.56	0.0396876	0.56	-1.59627
Extreme Value Sophisticate	0.57	0.0402888	0.57	-1.59638
Logistic Naive	0.76	0.00267137	0.76	-1.59188
Logistic Sophisticate	0.76	0.00331131	0.76	-1.59189
Laplace Naive	0.63	0.008065	0.63	-1.59202
Laplace Sophisticate	0.64	0.00933172	0.63	-1.59207

Table: Parameter estimates of β and squared distance and log-likelihood.

Parameter Estimates of β

Parameter Estimates of β

- Estimates vary widely depending on the functional form assumption.

Parameter Estimates of β

- Estimates vary widely depending on the functional form assumption.
- β is missestimated.

Parameter Estimates of β

- Estimates vary widely depending on the functional form assumption.
- β is missestimated.
- the example is robust to

Parameter Estimates of β

- Estimates vary widely depending on the functional form assumption.
- β is missestimated.
- the example is robust to
  - ▸ simultaneous estimation of mean and variance

Parameter Estimates of β

- Estimates vary widely depending on the functional form assumption.
- β is missestimated.
- the example is robust to
  - ▸ simultaneous estimation of mean and variance
  - more periods ▸ 30 periods ▸ 60 periods

Parameter Estimates of β

- Estimates vary widely depending on the functional form assumption.
- β is missestimated.
- the example is robust to
  - ▸ simultaneous estimation of mean and variance
  - more periods ▸ 30 periods ▸ 60 periods
  - More data can make the estimates worse ▸ example

Parameter Estimates of β

- Estimates vary widely depending on the functional form assumption.
- β is missestimated.
- the example is robust to
  - ▸ simultaneous estimation of mean and variance
  - more periods ▸ 30 periods ▸ 60 periods
  - More data can make the estimates worse ▸ example
  - ▸ different true distributions

Parameter Estimates of β

- Estimates vary widely depending on the functional form assumption.
- β is missestimated.
- the example is robust to
  - ▸ simultaneous estimation of mean and variance
  - more periods ▸ 30 periods ▸ 60 periods
  - More data can make the estimates worse ▸ example
  - ▸ different true distributions
  - ▸ a truely time-inconsistent agent $\beta = 0.9$

Parameter Estimates of β

- Estimates vary widely depending on the functional form assumption.
- β is missestimated.
- the example is robust to
  - ▸ simultaneous estimation of mean and variance
  - more periods ▸ 30 periods ▸ 60 periods
  - More data can make the estimates worse ▸ example
  - ▸ different true distributions
  - ▸ a truely time-inconsistent agent $\beta = 0.9$

Parameter Estimates of β

- Estimates vary widely depending on the functional form assumption.
- β is missestimated.
- the example is robust to
  - simultaneous estimation of mean and variance
  - more periods 30 periods 60 periods
  - More data can make the estimates worse example
  - different true distributions
  - a truely time-inconsistent agent $\beta = 0.9$
- Our theoretical results show that for **every dataset** estimates will be driven by functional form assumption.

Agent's Behavior

Preliminary Analysis

Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \underline{y}$.

Preliminary Analysis

Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \underline{y}$.

Earlier selves think Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \hat{\beta}\delta \frac{\underline{y}}{\beta\delta} = \left(\hat{\beta}/\beta\right) \underline{y}$.

Preliminary Analysis

Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \underline{y}$.

Earlier selves think Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \hat{\beta}\delta\frac{\underline{y}}{\beta\delta} = \left(\hat{\beta}/\beta\right)\underline{y}$.

By induction, all earlier selves (are perceived to) use cutoff strategies.

Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \underline{y}$.

Earlier selves think Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \hat{\beta}\delta \frac{\underline{y}}{\beta\delta} = \left(\hat{\beta}/\beta\right) \underline{y}$.

By induction, all earlier selves (are perceived to) use cutoff strategies.

- Let v_t be the perceived continuation value or **actual cutoff** of Self t.

Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \underline{y}$.

Earlier selves think Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \hat{\beta}\delta \frac{\underline{y}}{\beta\delta} = \left(\hat{\beta}/\beta\right) \underline{y}$.

By induction, all earlier selves (are perceived to) use cutoff strategies.

- Let v_t be the perceived continuation value or **actual cutoff** of Self t.
- Let $c_t = \left(\hat{\beta}/\beta\right) v_t$ be the cutoff earlier selves think t uses.

Preliminary Analysis

Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \underline{y}$.

Earlier selves think Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \hat{\beta}\delta \frac{\underline{y}}{\beta\delta} = \left(\hat{\beta}/\beta\right) \underline{y}$.

By induction, all earlier selves (are perceived to) use cutoff strategies.

- Let v_t be the perceived continuation value or **actual cutoff** of Self t.
- Let $c_t = \left(\hat{\beta}/\beta\right) v_t$ be the cutoff earlier selves think t uses.

For $t < T$, the perceived continuation values v_t satisfy the equation

$$v_t = \underbrace{\beta\,\delta \int_{c_{t+1}}^{\infty} z\, dF(z)}_{\text{anti. payoff of doing task tomorrow}} + \underbrace{F(c_{t+1})\,\delta\, v_{t+1}}_{\text{anti. payoff of continuing}} .$$

Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \underline{y}$.

Earlier selves think Self T uses a cutoff strategy.

- Completes the task if and only if $y_T > \hat{\beta}\delta\frac{\underline{y}}{\beta\delta} = \left(\hat{\beta}/\beta\right)\underline{y}$.

By induction, all earlier selves (are perceived to) use cutoff strategies.

- Let v_t be the perceived continuation value or **actual cutoff** of Self t.
- Let $c_t = \left(\hat{\beta}/\beta\right) v_t$ be the cutoff earlier selves think t uses.

For $t < T$, the perceived continuation values v_t satisfy the equation

$$v_t = \underbrace{\beta\,\delta\int_{c_{t+1}}^{\infty} z\,dF(z)}_{\text{anti. payoff of doing task tomorrow}} + \underbrace{F(c_{t+1})\,\delta\,v_{t+1}}_{\text{anti. payoff of continuing}}.$$

- Relative to tomorrow's self, discount the perceived continuation value by extra δ. $\Rightarrow$ simple recursive structure!

Rate of Task Completion Increases Over Time

Theorem (Monotonicity of the Continuation Value)

i.) The subjective continuation value is non-increasing over time

$$v_1 \geq v_2 \geq \ldots \geq v_T \, .$$

ii.) Every self t prefers a later deadline.

Rate of Task Completion Increases Over Time

Theorem (Monotonicity of the Continuation Value)

i.) The subjective continuation value is non-increasing over time

$$v_1 \geq v_2 \geq \ldots \geq v_T \, .$$

ii.) Every self t prefers a later deadline.

i.) $\Rightarrow$ ii.)

- In our stationary environment, perceived continuation value depends on the distance to the deadline.

Theorem (Monotonicity of the Continuation Value)

i.) The subjective continuation value is non-increasing over time

$$v_1 \geq v_2 \geq \ldots \geq v_T \, .$$

ii.) Every self t prefers a later deadline.

i.) $\Rightarrow$ ii.)

- In our stationary environment, perceived continuation value depends on the distance to the deadline.
- Since the shorter the distance, the lower the continuation value by i.), you never want to impose an earlier deadline when F is time-independent.

Theorem (Monotonicity of the Continuation Value)

i.) The subjective continuation value is non-increasing over time

$$v_1 \geq v_2 \geq \ldots \geq v_T \,.$$

ii.) Every self t prefers a later deadline.

i.) $\Rightarrow$ ii.)

- In our stationary environment, perceived continuation value depends on the distance to the deadline.
- Since the shorter the distance, the lower the continuation value by i.), you never want to impose an earlier deadline when F is time-independent.
  - Deadlines used to classify agents (Ariely and Wertenbroch, 2002; Bisin and Hyndman, 2018) as sophistcated time-inconsistent ones.

Rate of Task Completion Increases Over Time

Theorem (Monotonicity of the Continuation Value)

i.) The subjective continuation value is non-increasing over time

$$v_1 \geq v_2 \geq \ldots \geq v_T \, .$$

ii.) Every self t prefers a later deadline.

Corollary: the observed conditional stopping probability is non-decreasing toward the deadline; i.e.

$$p_1 \leq p_2 \leq \ldots \leq p_T \, .$$

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task is always costly:

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task is always costly:

- Consider a sophisticated hyperbolic discounter.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task is always costly:

- Consider a sophisticated hyperbolic discounter.
- Let Self 1 compare a $(T-1)$-period to T-period deadline.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task is always costly:

- Consider a sophisticated hyperbolic discounter.
- Let Self 1 compare a $(T-1)$-period to T-period deadline.
- If Self 1 waits in the T-period problem, Self 2 faces a $(T-1)$-period problem.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task is always costly:

- Consider a sophisticated hyperbolic discounter.
- Let Self 1 compare a $(T-1)$-period to T-period deadline.
- If Self 1 waits in the T-period problem, Self 2 faces a $(T-1)$-period problem.
- Future selves behavior s-periods before the deadline is identical, and so is task completion s periods before the deadline.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task is always costly:

- Consider a sophisticated hyperbolic discounter.
- Let Self 1 compare a $(T-1)$-period to T-period deadline.
- If Self 1 waits in the T-period problem, Self 2 faces a $(T-1)$-period problem.
- Future selves behavior s-periods before the deadline is identical, and so is task completion s periods before the deadline.
- Due to discounting, Self 1 is strictly better off selecting the T-period problem and not doing the task in the first period.

Since partially naive agents think they are sophisticated, and soph. agents never benefits, they also do not impose a deadline.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task always yields a positive payoff:

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task always yields a positive payoff:

- Future selves are to impatient, and therefore too willing to stop.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task always yields a positive payoff:

- Future selves are to impatient, and therefore too willing to stop.
- Suppose Self 1 can extend the deadline from $T-1$ to T periods.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task always yields a positive payoff:

- Future selves are to impatient, and therefore too willing to stop.
- Suppose Self 1 can extend the deadline from $T-1$ to T periods.
- With a longer deadline, Self $T-1$ will wait for sufficiently low net benefits.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task always yields a positive payoff:

- Future selves are to impatient, and therefore too willing to stop.
- Suppose Self 1 can extend the deadline from $T-1$ to T periods.
- With a longer deadline, Self $T-1$ will wait for sufficiently low net benefits.
- Whenever the impatient Self $T-1$ chooses to wait, Self 1 strongly prefers it to wait.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task always yields a positive payoff:

- Future selves are to impatient, and therefore too willing to stop.
- Suppose Self 1 can extend the deadline from $T-1$ to T periods.
- With a longer deadline, Self $T-1$ will wait for sufficiently low net benefits.
- Whenever the impatient Self $T-1$ chooses to wait, Self 1 strongly prefers it to wait.
- Conditional on reaching period $T-1$, thus, the longer deadline benefits Self 1.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task always yields a positive payoff:

- Future selves are to impatient, and therefore too willing to stop.
- Suppose Self 1 can extend the deadline from $T-1$ to T periods.
- With a longer deadline, Self $T-1$ will wait for sufficiently low net benefits.
- Whenever the impatient Self $T-1$ chooses to wait, Self 1 strongly prefers it to wait.
- Conditional on reaching period $T-1$, thus, the longer deadline benefits Self 1.
- With the longer deadline, Self $T-2$'s benefits from waiting increases as it also thinks of $T-1$ as to impatient.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task always yields a positive payoff:

- Future selves are to impatient, and therefore too willing to stop.
- Suppose Self 1 can extend the deadline from $T-1$ to T periods.
- With a longer deadline, Self $T-1$ will wait for sufficiently low net benefits.
- Whenever the impatient Self $T-1$ chooses to wait, Self 1 strongly prefers it to wait.
- Conditional on reaching period $T-1$, thus, the longer deadline benefits Self 1.
- With the longer deadline, Self $T-2$'s benefits from waiting increases as it also thinks of $T-1$ as to impatient.
- Hence, Self $T-2$ will also act less impatiently, which again benefits Self 1 conditional on reaching period $T-2$.

Why is the Perceived Continuation Value Decreasing?

Intuition when doing the task always yields a positive payoff:

- Future selves are to impatient, and therefore too willing to stop.
- Suppose Self 1 can extend the deadline from $T-1$ to T periods.
- With a longer deadline, Self $T-1$ will wait for sufficiently low net benefits.
- Whenever the impatient Self $T-1$ chooses to wait, Self 1 strongly prefers it to wait.
- Conditional on reaching period $T-1$, thus, the longer deadline benefits Self 1.
- With the longer deadline, Self $T-2$'s benefits from waiting increases as it also thinks of $T-1$ as to impatient.
- Hence, Self $T-2$ will also act less impatiently, which again benefits Self 1 conditional on reaching period $T-2$.
- .... Self 1 benefits from longer deadline.

Non-Identifiability

Theorem (Non-identifiability for Fully Naive Case: $\hat{\beta} = 1$)

For every non-decreasing sequence of stopping probabilities $0 < p_1 \leq p_2 \leq \ldots \leq p_T < 1$, *every* $(\delta, \beta) \in (0,1) \times (0,1]$, *and every penalty* $\underline{y}/\beta\delta < 0$, *there exists a distribution* F *that rationalizes the agent's stopping probabilities as the unique outcome of any perception perfect equilibrium.*

Theorem (Non-identifiability for Sophisticated Case: $\hat{\beta} = \beta$)

For every non-decreasing sequence of stopping probabilities $0 < p_1 \leq p_2 \leq \ldots \leq p_T < 1$, *every* $(\delta, \beta) \in (0,1] \times (0,1]$, *and every penalty* $\underline{y}/\beta\delta$, *there exists a distribution* F *that rationalizes the agent's stopping probabilities as the outcome of a perception perfect equilibrium.*

Why are Time-Preferences Unidentifiable?

Rough Intuition:

- Whether a self prefers to do a task today or wait depends on her time preferences and on the perceived option value of waiting.
- The option value of waiting depends on the payoff distribution.
- By changing the unobservable payoff distribution, can undo a change in the time preference of the agent.

Why are Time-Preferences Unidentifiable?

Rough Intuition:

- Whether a self prefers to do a task today or wait depends on her time preferences and on the perceived option value of waiting.
- The option value of waiting depends on the payoff distribution.
- By changing the unobservable payoff distribution, can undo a change in the time preference of the agent.
- But, since a local change in the payoff distribution affects all continuation values in a highly non-linear way, need a non-local argument.

Why are Time-Preferences Unidentifiable?

Rough Intuition:

- Whether a self prefers to do a task today or wait depends on her time preferences and on the perceived option value of waiting.
- The option value of waiting depends on the payoff distribution.
- By changing the unobservable payoff distribution, can undo a change in the time preference of the agent.
- But, since a local change in the payoff distribution affects all continuation values in a highly non-linear way, need a non-local argument.

For non-local argument:

- Fully naive case: map continuation values into payoff distributions and back to continuation values in a "monotone way" that allows using Tarsky's Theorem.

Why are Time-Preferences Unidentifiable?

Rough Intuition:

- Whether a self prefers to do a task today or wait depends on her time preferences and on the perceived option value of waiting.
- The option value of waiting depends on the payoff distribution.
- By changing the unobservable payoff distribution, can undo a change in the time preference of the agent.
- But, since a local change in the payoff distribution affects all continuation values in a highly non-linear way, need a non-local argument.

For non-local argument:

- Fully naive case: map continuation values into payoff distributions and back to continuation values in a "monotone way" that allows using Tarsky's Theorem.
- Sophisticated case: focus on distributions for which the recursive structure for continuation values gives rise to a linear system of equations (which can be solved forward).

Jump to A Priori Knowledge.

Non-Identifiability for Naives

Outline of the Proof.

1. Fix time preferences and $\underline{y}$.

Outline of the Proof.

1. Fix time preferences and $\underline{y}$.
2. Take a vector of perc. continuation values $v_1 \geq v_2 \geq \ldots \geq v_{T-1}$.

Outline of the Proof.

1. Fix time preferences and $\underline{y}$.
2. Take a vector of perc. continuation values $v_1 \geq v_2 \geq \ldots \geq v_{T-1}$.
3. Generate a payoff distribution that gives the desired stopping probabilities.
   - Put probability mass equal to $p_t - p_{t-1}$ between v_t and v_{t-1}.
   - Maps continuation values into distributions with correct stopping prob.

Outline of the Proof.

1. Fix time preferences and $\underline{y}$.
2. Take a vector of perc. continuation values $v_1 \geq v_2 \geq \ldots \geq v_{T-1}$.
3. Generate a payoff distribution that gives the desired stopping probabilities.
   - Put probability mass equal to $p_t - p_{t-1}$ between v_t and v_{t-1}.
   - Maps continuation values into distributions with correct stopping prob.
4. Calculate the actual continuation values.
   - Maps the set of distributions back into the vector of continuation values.
   - By earlier theorem, these continuation values are non-increasing.
   - So function maps a non-increasing sequence into a non-increasing sequence.

Outline of the Proof.

1. Fix time preferences and $\underline{y}$.
2. Take a vector of perc. continuation values $v_1 \geq v_2 \geq \ldots \geq v_{T-1}$.
3. Generate a payoff distribution that gives the desired stopping probabilities.
   - Put probability mass equal to $p_t - p_{t-1}$ between v_t and v_{t-1}.
   - Maps continuation values into distributions with correct stopping prob.
4. Calculate the actual continuation values.
   - Maps the set of distributions back into the vector of continuation values.
   - By earlier theorem, these continuation values are non-increasing.
   - So function maps a non-increasing sequence into a non-increasing sequence.
5. Function is bounded, maps sequences from an appropriately chosen interval into itself, and monotone.
   - Higher continuation values lead to a better distribution (in the sense of first-order stochastic dominance) and for $\hat{\beta} = 1$ a better distribution increases the subjective continuation values.

Outline of the Proof.

1. Fix time preferences and $\underline{y}$.
2. Take a vector of perc. continuation values $v_1 \geq v_2 \geq \ldots \geq v_{T-1}$.
3. Generate a payoff distribution that gives the desired stopping probabilities.
   - Put probability mass equal to $p_t - p_{t-1}$ between v_t and v_{t-1}.
   - Maps continuation values into distributions with correct stopping prob.
4. Calculate the actual continuation values.
   - Maps the set of distributions back into the vector of continuation values.
   - By earlier theorem, these continuation values are non-increasing.
   - So function maps a non-increasing sequence into a non-increasing sequence.
5. Function is bounded, maps sequences from an appropriately chosen interval into itself, and monotone.
   - Higher continuation values lead to a better distribution (in the sense of first-order stochastic dominance) and for $\hat{\beta} = 1$ a better distribution increases the subjective continuation values.
6. Apply Tarski's Theorem.

Monotonicity in Payoffs and Naivete

Important step: a FOSD increase in F raises continuation value (with $\hat{\beta} = 1$).

Monotonicity in Payoffs and Naivete

Important step: a FOSD increase in F raises continuation value (with $\hat{\beta} = 1$).

- For a time-consistent agent, a FOSD in F increases v_{T-1}.

Monotonicity in Payoffs and Naivete

Important step: a FOSD increase in F raises continuation value (with $\hat{\beta} = 1$).

- For a time-consistent agent, a FOSD in F increases v_{T-1}.
- From perspective of $T-2$, for a time-consistent agent, above plus the FOSD makes it more desirable to reach $T-1$.

Monotonicity in Payoffs and Naivete

Important step: a FOSD increase in F raises continuation value (with $\hat{\beta} = 1$).

- For a time-consistent agent, a FOSD in F increases v_{T-1}.
- From perspective of $T-2$, for a time-consistent agent, above plus the FOSD makes it more desirable to reach $T-1$.
- Since agent thinks she acts time-consistently from tomorrow on, v_t's increase.

Monotonicity in Payoffs and Naivete

Important step: a FOSD increase in F raises continuation value (with $\hat{\beta} = 1$).

- For a time-consistent agent, a FOSD in F increases v_{T-1}.
- From perspective of $T - 2$, for a time-consistent agent, above plus the FOSD makes it more desirable to reach $T - 1$.
- Since agent thinks she acts time-consistently from tomorrow on, v_t's increase.

We show via a simple example that a FOSD in F need not raise continuation values if $\hat{\beta} < 1$.

- Ex.: soph. agent prefers to pay lump-sum tax when doing task.
- $T = 3$. Payoffs are either high or low.

Monotonicity in Payoffs and Naivete

Important step: a FOSD increase in F raises continuation value (with $\hat{\beta} = 1$).

- For a time-consistent agent, a FOSD in F increases v_{T-1}.
- From perspective of $T - 2$, for a time-consistent agent, above plus the FOSD makes it more desirable to reach $T - 1$.
- Since agent thinks she acts time-consistently from tomorrow on, v_t's increase.

We show via a simple example that a FOSD in F need not raise continuation values if $\hat{\beta} < 1$.

- Ex.: soph. agent prefers to pay lump-sum tax when doing task.
- $T = 3$. Payoffs are either high or low.
- Tax eliminates the temptation to quit in period 2 for low payoff realization.

Monotonicity in Payoffs and Naivete

Important step: a FOSD increase in F raises continuation value (with $\hat{\beta} = 1$).

- For a time-consistent agent, a FOSD in F increases v_{T-1}.
- From perspective of $T - 2$, for a time-consistent agent, above plus the FOSD makes it more desirable to reach $T - 1$.
- Since agent thinks she acts time-consistently from tomorrow on, v_t's increase.

We show via a simple example that a FOSD in F need not raise continuation values if $\hat{\beta} < 1$.

- Ex.: soph. agent prefers to pay lump-sum tax when doing task.
- $T = 3$. Payoffs are either high or low.
- Tax eliminates the temptation to quit in period 2 for low payoff realization.
- Commitment to only stop when payoff are high in $t = 2$ increases v_1.

Monotonicity in Payoffs and Naivete

Important step: a FOSD increase in F raises continuation value (with $\hat{\beta} = 1$).

- For a time-consistent agent, a FOSD in F increases v_{T-1}.
- From perspective of $T-2$, for a time-consistent agent, above plus the FOSD makes it more desirable to reach $T-1$.
- Since agent thinks she acts time-consistently from tomorrow on, v_t's increase.

We show via a simple example that a FOSD in F need not raise continuation values if $\hat{\beta} < 1$.

- Ex.: soph. agent prefers to pay lump-sum tax when doing task.
- $T = 3$. Payoffs are either high or low.
- Tax eliminates the temptation to quit in period 2 for low payoff realization.
- Commitment to only stop when payoff are high in $t = 2$ increases v_1.
- Benefits of commitment overcompensate the direct payoff reduction through the tax.

Non-Identifiability for Sophisticates

Time-Preferences are Unidentifiable: Sophisticated Case

Theorem (Non-identifiability for $\hat{\beta} = \beta$)

For every non-decreasing sequence of stopping probabilities $0 < p_1 \leq p_2 \leq \ldots \leq p_T < 1$, every $(\delta, \beta) \in (0,1] \times (0,1]$, and every penalty $\underline{y}/\beta\delta$, there exists a distribution F that rationalizes the agent's stopping probabilities as the outcome of a perception perfect equilibrium.

Time-Preferences are Unidentifiable: Sophisticated Case

Theorem (Non-identifiability for $\hat{\beta} = \beta$)

For every non-decreasing sequence of stopping probabilities $0 < p_1 \leq p_2 \leq \ldots \leq p_T < 1$, *every* $(\delta, \beta) \in (0,1] \times (0,1]$, *and every penalty* $\underline{y}/\beta\delta$, *there exists a distribution* F *that rationalizes the agent's stopping probabilities as the outcome of a perception perfect equilibrium.*

Rough outline of the proof:

Time-Preferences are Unidentifiable: Sophisticated Case

Theorem (Non-identifiability for $\hat{\beta} = \beta$)

For every non-decreasing sequence of stopping probabilities $0 < p_1 \leq p_2 \leq \ldots \leq p_T < 1$, every $(\delta, \beta) \in (0,1] \times (0,1]$, and every penalty $\underline{y}/\beta\delta$, there exists a distribution F that rationalizes the agent's stopping probabilities as the outcome of a perception perfect equilibrium.

Rough outline of the proof:

- Choose a distribution with $t+1$ interals with constant density, of which the non-extreme values are set equal to the continuation values. (With the second lowest value being set at $v_T = \underline{y}$, etc... .)

Theorem (Non-identifiability for $\hat{\beta} = \beta$)

For every non-decreasing sequence of stopping probabilities $0 < p_1 \leq p_2 \leq \ldots \leq p_T < 1$, *every* $(\delta, \beta) \in (0,1] \times (0,1]$, *and every penalty* $\underline{y}/\beta\delta$, *there exists a distribution* F *that rationalizes the agent's stopping probabilities as the outcome of a perception perfect equilibrium.*

Rough outline of the proof:

- Choose a distribution with $t+1$ interals with constant density, of which the non-extreme values are set equal to the continuation values. (With the second lowest value being set at $v_T = \underline{y}$, etc... .)
- Select the probability mass on the intervals to match the increasing stopping probabilities.

Time-Preferences are Unidentifiable: Sophisticated Case

Theorem (Non-identifiability for $\hat{\beta} = \beta$)

For every non-decreasing sequence of stopping probabilities $0 < p_1 \leq p_2 \leq \ldots \leq p_T < 1$, every $(\delta, \beta) \in (0,1] \times (0,1]$, and every penalty $\underline{y}/\beta\delta$, there exists a distribution F that rationalizes the agent's stopping probabilities as the outcome of a perception perfect equilibrium.

Rough outline of the proof:

- Choose a distribution with $t + 1$ interals with constant density, of which the non-extreme values are set equal to the continuation values. (With the second lowest value being set at $v_T = \underline{y}$, etc... .)
- Select the probability mass on the intervals to match the increasing stopping probabilities.
- When $\hat{\beta} = \beta$, the recursive structure for continuation values in this case gives rise to a linear system of equations.

Theorem (Non-identifiability for $\hat{\beta} = \beta$)

For every non-decreasing sequence of stopping probabilities $0 < p_1 \leq p_2 \leq \ldots \leq p_T < 1$, every $(\delta, \beta) \in (0,1] \times (0,1]$, and every penalty $\underline{y}/\beta\delta$, there exists a distribution F that rationalizes the agent's stopping probabilities as the outcome of a perception perfect equilibrium.

Rough outline of the proof:

- Choose a distribution with $t+1$ interals with constant density, of which the non-extreme values are set equal to the continuation values. (With the second lowest value being set at $v_T = \underline{y}$, etc... .)
- Select the probability mass on the intervals to match the increasing stopping probabilities.
- When $\hat{\beta} = \beta$, the recursive structure for continuation values in this case gives rise to a linear system of equations.
- Can solve forward for all continuation values, and if lowest mass point is low enough, gives rise to well-defined solution.

A Priori Knowledge

Knowledge of Expected Value and Variance (but not Penalty)

Consider the case of a time-consistent agent that is fully patient $\beta = \delta = 1$.

Knowledge of Expected Value and Variance (but not Penalty)

Consider the case of a time-consistent agent that is fully patient $\beta = \delta = 1$.

- By our last theorem, we can rationalize the agent's stopping behavior through some distribution F.

Knowledge of Expected Value and Variance (but not Penalty)

Consider the case of a time-consistent agent that is fully patient $\beta = \delta = 1$.

- By our last theorem, we can rationalize the agent's stopping behavior through some distribution F.
- For $\beta = \delta = 1$, we can determine the continuation value recursively as

$$v_t = \mathbb{E}\left[\max\{y_{t+1}, v_{t+1}\}\right].$$

Knowledge of Expected Value and Variance (but not Penalty)

Consider the case of a time-consistent agent that is fully patient $\beta = \delta = 1$.

- By our last theorem, we can rationalize the agent's stopping behavior through some distribution F.
- For $\beta = \delta = 1$, we can determine the continuation value recursively as

$$v_t = \mathbb{E}\left[\max\{y_{t+1}, v_{t+1}\}\right].$$

- Hence for any $c_1, c_2 > 0$, the stopping behavior is also optimal for the payoff $c_1 y + c_2$ and penalty $c_1 \underline{y} + c_2$.

Knowledge of Expected Value and Variance (but not Penalty)

Consider the case of a time-consistent agent that is fully patient $\beta = \delta = 1$.

- By our last theorem, we can rationalize the agent's stopping behavior through some distribution F.
- For $\beta = \delta = 1$, we can determine the continuation value recursively as

$$v_t = \mathbb{E}\left[\max\{y_{t+1}, v_{t+1}\}\right].$$

- Hence for any $c_1, c_2 > 0$, the stopping behavior is also optimal for the payoff $c_1 y + c_2$ and penalty $c_1 \underline{y} + c_2$.
- Thus, we can freely choose mean and variance of F and still match the observed stopping behavior.

Knowledge of Expected Value and Variance (but not Penalty)

Consider the case of a time-consistent agent that is fully patient $\beta = \delta = 1$.

- By our last theorem, we can rationalize the agent's stopping behavior through some distribution F.
- For $\beta = \delta = 1$, we can determine the continuation value recursively as

$$v_t = \mathbb{E}\left[\max\{y_{t+1}, v_{t+1}\}\right].$$

- Hence for any $c_1, c_2 > 0$, the stopping behavior is also optimal for the payoff $c_1 y + c_2$ and penalty $c_1 \underline{y} + c_2$.
- Thus, we can freely choose mean and variance of F and still match the observed stopping behavior.
- Parameteric identification of β must rely on other features of the distribution!

Jump to Conclusion.

Rich Data

Observable Perceived Continuation Values

Analyst's problem: disentangle **desire to delay** from **option value**:

Observable Perceived Continuation Values

Analyst's problem: disentangle **desire to delay** from **option value**:

- High option value requires payoff to differ significantly.

Observable Perceived Continuation Values

Analyst's problem: disentangle **desire to delay** from **option value**:

- High option value requires payoff to differ significantly.
- But then perc. continuation value should drop quickly as deadline approaches.

Observable Perceived Continuation Values

Analyst's problem: disentangle **desire to delay** from **option value**:

- High option value requires payoff to differ significantly.
- But then perc. continuation value should drop quickly as deadline approaches.
- Present-bias can lead to delay even if option value is (relatively) constant.

Observable Perceived Continuation Values

Analyst's problem: disentangle **desire to delay** from **option value**:

- High option value requires payoff to differ significantly.
- But then perc. continuation value should drop quickly as deadline approaches.
- Present-bias can lead to delay even if option value is (relatively) constant.

Can observing option value help with non-parametric identification?

Observable Perceived Continuation Values

Analyst's problem: disentangle **desire to delay** from **option value**:

- High option value requires payoff to differ significantly.
- But then perc. continuation value should drop quickly as deadline approaches.
- Present-bias can lead to delay even if option value is (relatively) constant.

Can observing option value help with non-parametric identification?

- For clean answer, suppose contemporaneous utility is linear in money and agent sophisticated.

Observable Perceived Continuation Values

Analyst's problem: disentangle **desire to delay** from **option value**:

- High option value requires payoff to differ significantly.
- But then perc. continuation value should drop quickly as deadline approaches.
- Present-bias can lead to delay even if option value is (relatively) constant.

Can observing option value help with non-parametric identification?

- For clean answer, suppose contemporaneous utility is linear in money and agent sophisticated.
- Aside: since need to ask only once, analyst does not (implicitly) elicit time-preferences over money (see Ericson and Laibson, 2019; Ramsey, 1928, for why this is important).

Simplifying from Infinite-Dimensional to Finite Dimensional Space

- Intuitively, option value of waiting is determined by
  - probability with which agents stops at given t;
  - expected payoff conditional on stopping at t.

Simplifying from Infinite-Dimensional to Finite Dimensional Space

- Intuitively, option value of waiting is determined by
  - probability with which agents stops at given t;
  - expected payoff conditional on stopping at t.
- Can move probability mass between any two perc. continuation values to expected payoff conditional on falling between these.
  - Does not change continuation values or stopping probabilities.

Simplifying from Infinite-Dimensional to Finite Dimensional Space

- Intuitively, option value of waiting is determined by
  - probability with which agents stops at given t;
  - expected payoff conditional on stopping at t.
- Can move probability mass between any two perc. continuation values to expected payoff conditional on falling between these.
  - Does not change continuation values or stopping probabilities.
- Analyst only needs to consider distributions with $T+1$ mass points falling between the continuation values.

Simplifying from Infinite-Dimensional to Finite Dimensional Space

- Intuitively, option value of waiting is determined by
  - probability with which agents stops at given t;
  - expected payoff conditional on stopping at t.
- Can move probability mass between any two perc. continuation values to expected payoff conditional on falling between these.
  - Does not change continuation values or stopping probabilities.
- Analyst only needs to consider distributions with $T+1$ mass points falling between the continuation values.

We say the data is *plausible* if conditional stopping probabilities are increasing and continuation values decreasing.

Theorem

Suppose $u(m_t) = m_t$ for all t and that $p_1 > 0$. Plausible data (v, p) is consistent with β, δ and sophistication $\hat{\beta} = \beta$ if and only if (i)

$$\beta < \frac{\delta^{-1} v_1 - (1 - p_2) v_2}{v_2(p_2 - p_1) + v_1 p_1}$$

and (ii) $v_{t+1}\beta < v_{t+1} a(\delta, t) \leq v_t \beta$ for all $t \in \{2, \ldots, T-1\}$, where

$$a(\delta, t) = 1 - \frac{\delta^{-1}(v_{t-1} - v_t) - (1 - p_t)(v_t - v_{t+1})}{v_{t+1}(p_{t+1} - p_t)}.$$

Boils down to checking a simple set of inequalities.

Theorem

Suppose $u(m_t) = m_t$ for all t and that $p_1 > 0$. Plausible data (v, p) is consistent with β, δ and sophistication $\hat{\beta} = \beta$ if and only if (i)

$$\beta < \frac{\delta^{-1} v_1 - (1 - p_2) v_2}{v_2(p_2 - p_1) + v_1 p_1}$$

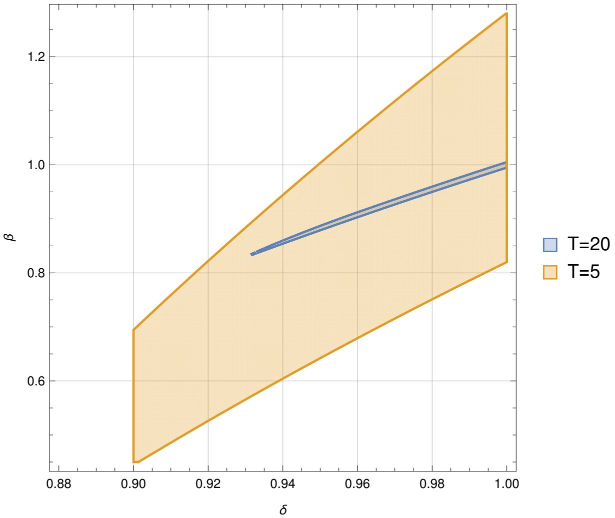
and (ii) $v_{t+1}\beta < v_{t+1}a(\delta, t) \leq v_t\beta$ for all $t \in \{2, \ldots, T-1\}$, where

$$a(\delta, t) = 1 - \frac{\delta^{-1}(v_{t-1} - v_t) - (1 - p_t)(v_t - v_{t+1})}{v_{t+1}(p_{t+1} - p_t)}.$$

Boils down to checking a simple set of inequalities.

Extends to non-linear utility and partial naivete at cost of using numerical techniques.

Consistent Parameter Estimates for Example with $T = 5$ and $T = 20$



Jump to Conclusion.

Relationship to Dynamic Discrete Choice Literature

Conceptual Difference to Dynamic Discrete Choice Literature

Vast literature on dynamic discrete choice considers identification of

- time preferences; and
- instantaneous payoffs.

Dynamic Discrete Choice focusses on:

1. non-parametric state and action dependent mean utility (state = time $\Rightarrow$ non-iid data);
2. unobservable shock is distributed with some known distribution (e.g., extreme-value type 1).

Conceptual Difference to Dynamic Discrete Choice Literature

Vast literature on dynamic discrete choice considers identification of

- time preferences; and
- instantaneous payoffs.

Dynamic Discrete Choice focusses on:

1. non-parametric state and action dependent mean utility (state = time $\Rightarrow$ non-iid data);
2. unobservable shock is distributed with some known distribution (e.g., extreme-value type 1).

We focus on:

1. single unknown mean utility level;
2. non-parametric in the distribution of the unobservable shock.

Common Setup in Dynamic Discrete Choice Literature

Common setup in dynamic discrete choice literature:

1. infinite horizon;
2. agent is time-consistent;
3. feasible actions do not depend on past actions;
4. additive separability between observable part and shock; and
5. shocks are drawn from some (typically given) distribution with unbounded support.

Formally, 3. to 5. rule out stopping problems.

Common Setup in Dynamic Discrete Choice Literature

Common setup in dynamic discrete choice literature:

1. infinite horizon;
2. agent is time-consistent;
3. feasible actions do not depend on past actions;
4. additive separability between observable part and shock; and
5. shocks are drawn from some (typically given) distribution with unbounded support.

Formally, 3. to 5. rule out stopping problems.

Common Setup in Dynamic Discrete Choice Literature

Common setup in dynamic discrete choice literature:

1. infinite horizon;
2. agent is time-consistent;
3. feasible actions do not depend on past actions;
4. additive separability between observable part and shock; and
5. shocks are drawn from some (typically given) distribution with unbounded support.

Formally, 3. to 5. rule out stopping problems.

- Nevertheless, our results "question" some existing parametric identification ideas.

Non-Identification in Dynamic Discrete Choice Literature

Classic **parametric** non-identification result (e.g., Section 3.5 in Rust, 1994; Magnac and Thesmar, 2002) of dynamic discrete choice literature:

- With a state-dependent shock (or mean utility), for any known invertible distribution of unobservable payoffs impossible to identify time-preference parameter.
- Corresponding state in our setting is time to deadline.
- Result extends straightforwardly to our setting for any combination of $(\delta, \beta, \hat{\beta})$.

Non-Identification in Dynamic Discrete Choice Literature

Classic **parametric** non-identification result (e.g., Section 3.5 in Rust, 1994; Magnac and Thesmar, 2002) of dynamic discrete choice literature:

- With a state-dependent shock (or mean utility), for any known invertible distribution of unobservable payoffs impossible to identify time-preference parameter.
- Corresponding state in our setting is time to deadline.
- Result extends straightforwardly to our setting for any combination of $(\delta, \beta, \hat{\beta})$.

If unknown payoffs are iid, however, **parametric identification** possible for time-consistent agent (and beyond)!

Classic **parametric** non-identification result (e.g., Section 3.5 in Rust, 1994; Magnac and Thesmar, 2002) of dynamic discrete choice literature:

- With a state-dependent shock (or mean utility), for any known invertible distribution of unobservable payoffs impossible to identify time-preference parameter.
- Corresponding state in our setting is time to deadline.
- Result extends straightforwardly to our setting for any combination of $(\delta, \beta, \hat{\beta})$.

If unknown payoffs are iid, however, **parametric identification** possible for time-consistent agent (and beyond)!

Non-Identification in Dynamic Discrete Choice Literature

Classic **parametric** non-identification result (e.g., Section 3.5 in Rust, 1994; Magnac and Thesmar, 2002) of dynamic discrete choice literature:

- With a state-dependent shock (or mean utility), for any known invertible distribution of unobservable payoffs impossible to identify time-preference parameter.
- Corresponding state in our setting is time to deadline.
- Result extends straightforwardly to our setting for any combination of $(\delta, \beta, \hat{\beta})$.

If unknown payoffs are iid, however, **parametric identification** possible for time-consistent agent (and beyond)!

- Martinez et al. (2017) prove that β is identified when $\hat{\beta} = 1$, the analyst knows δ, and shocks are logistic with know variance.

As Levy and Schiraldi (2020)—who provide parametric identification results in β, δ dynamic discrete choice model with at least four actions—put it:

> *[a] typical approach to identification in the exponential discounting model adds exclusion restrictions on utility (conditional value function) across states, the presence of an absorbing choice (e.g. Magnac and Thesmar, 2002; Abbring and Daljord, 2019b), or restricts attention to a finite horizon model (e.g. Yao et al., 2012; Chung et al., 2014; Bajari et al., 2016; Chou, 2016), usually coupled with a strong normalization on the utility of the reference alternative.*

We imposes *all* of the above restrictions but our analyst doesn't know the parametric form of the distribution of shocks.

Related Approaches

Related Approaches

- Norets and Tang (2014) provide a system of equations for ("common") dynamic discrete binary choice environments that allows one to check (numerically) for a given δ whether it possible to find a stationary error distribution F that rationalizes the data.
  - No non-identification result in their environment.
  - Relates to our exercise with observable continuation values.

Related Approaches

- Norets and Tang (2014) provide a system of equations for ("common") dynamic discrete binary choice environments that allows one to check (numerically) for a given δ whether it possible to find a stationary error distribution F that rationalizes the data.
  - No non-identification result in their environment.
  - Relates to our exercise with observable continuation values.

- Christensen and Connault (2019) observe that predictions might depend on untestable assumptions on unobservable shocks and provide methods for "robust estimation"

Related Approaches

- Norets and Tang (2014) provide a system of equations for ("common") dynamic discrete binary choice environments that allows one to check (numerically) for a given δ whether it possible to find a stationary error distribution F that rationalizes the data.
  - No non-identification result in their environment.
  - Relates to our exercise with observable continuation values.

- Christensen and Connault (2019) observe that predictions might depend on untestable assumptions on unobservable shocks and provide methods for "robust estimation"
  - provide an dynamic discrete choice example to illustrate the crucial role of parametric assumptions.

Related Approaches

- Norets and Tang (2014) provide a system of equations for ("common") dynamic discrete binary choice environments that allows one to check (numerically) for a given δ whether it possible to find a stationary error distribution F that rationalizes the data.
  - No non-identification result in their environment.
  - Relates to our exercise with observable continuation values.
- Christensen and Connault (2019) observe that predictions might depend on untestable assumptions on unobservable shocks and provide methods for "robust estimation"
  - provide an dynamic discrete choice example to illustrate the crucial role of parametric assumptions.
  - Our result: in task-completion estimates are always driven by the paramteric assumption.

Related Approaches

- Norets and Tang (2014) provide a system of equations for ("common") dynamic discrete binary choice environments that allows one to check (numerically) for a given δ whether it possible to find a stationary error distribution F that rationalizes the data.
  - No non-identification result in their environment.
  - Relates to our exercise with observable continuation values.

- Christensen and Connault (2019) observe that predictions might depend on untestable assumptions on unobservable shocks and provide methods for "robust estimation"
  - provide an dynamic discrete choice example to illustrate the crucial role of parametric assumptions.
  - Our result: in task-completion estimates are always driven by the paramteric assumption.

- Imposing time-consistency, De Oliveira and Lamba (2019) characterize what an analyst can infer about δ when she observes an agent who chooses actions over time.
  - General decision environment.
  - A single sequence of actions instead of distribution.

Conclusion

- Homogeneity is important for predicting increasing stopping probability.

Conclusion

- Homogeneity is important for predicting increasing stopping probability.
  - Two groups with different (increasing) stopping probability suffice to generate non-monotonicity.

Conclusion

- Homogeneity is important for predicting increasing stopping probability.
- Even when knowing the first two moments of F, can always rationalize data if size of penalty unknown or task mandatory.

Conclusion

- Homogeneity is important for predicting increasing stopping probability.
- Even when knowing the first two moments of F, can always rationalize data if size of penalty unknown or task mandatory.
  - In that sense need "strong" parametric knowledge to do so.
  - Different F can rationalize data for same $\beta, \hat{\beta}, \delta$.

Conclusion

- Homogeneity is important for predicting increasing stopping probability.
- Even when knowing the first two moments of F, can always rationalize data if size of penalty unknown or task mandatory.
- "Proof of concept" for non-parameteric identification with rich data.

Conclusion

- Homogeneity is important for predicting increasing stopping probability.
- Even when knowing the first two moments of F, can always rationalize data if size of penalty unknown or task mandatory.
- "Proof of concept" for non-parameteric identification with rich data.
- Cannot infer time-preferences from bunching at the deadline even when having individual data.

Conclusion

- Homogeneity is important for predicting increasing stopping probability.
- Even when knowing the first two moments of F, can always rationalize data if size of penalty unknown or task mandatory.
- "Proof of concept" for non-parameteric identification with rich data.
- Cannot infer time-preferences from bunching at the deadline even when having individual data.
- Even sophisticated agents do not choose deadlines in stationary task-completion problem.
  - So no puzzle that people do not commit (in this environment).
- Most important: time-inconsistency may still be a major driver for why some agents complete tasks last minute.

Thank You!

References I

Akerlof, G. A. (1991). Procrastination and obedience. *American Economic Review*, 81(2):1–19.

Ariely, D. and Wertenbroch, K. (2002). Procrastination, deadlines, and preformance: Self-control by precommitment. *Psychological Science*, 13(3):219–224.

Augenblick, N., Niederle, M., and Sprenger, C. (2015). Working over time: Dynamic inconsistency in real effort tasks. *The Quarterly Journal of Economics*, 130(3):1067–1115.

Augenblick, N. and Rabin, M. (2019). An Experiment on Time Preference and Misprediction in Unpleasant Tasks. *Review of Economic Studies*, 86(3):941–975.

Bisin, A. and Hyndman, K. (2018). Present-bias, procrastination and deadlines in a field experiment. Working Paper.

Brown, J. R. and Previtero, A. (2018). Saving for retirement, annuities and procrastination. Working Paper.

Christensen, T. and Connault, B. (2019). Counterfactual sensitivity and robustness. *arXiv preprint arXiv:1904.00989*.

De Oliveira, H. and Lamba, R. (2019). Rationalizing dynamic choices. *Available at SSRN 3332092*.

Dixit, A. K. and Pindyck, R. S. (1994). *Investment under Uncertainty*. Princeton, NJ: Princeton University Press.

Ericson, K. M. M. and Laibson, D. I. (2019). Intertemporal choice. In Bernheim, B. D., DellaVigna, S., and Laibson, D. I., editors, *Handbook of Behavioral Economics - Foundations and Applications 2*. Elsevier.

Frakes, M. D. and Wasserman, M. F. (2016). Procrastination in the workplace: Evidence from the u.s. patent office. NBER Working Paper #22987.

Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40(2):351–401.

Heffetz, O., O'Donoghue, T., and Schneider, H. S. (2016). Forgetting and heterogeneity in task delay: Evidence from new york city parking-ticket recepients. NBER Working Paper #23012.

Magnac, T. and Thesmar, D. (2002). Identifying dynamic discrete decision processes. *Econometrica*, 70(2):801–816.

Martinez, S.-K., Meier, S., and Sprenger, C. (2017). Procrastination in the field: Evidence from tax filing. Working Paper.

O'Donoghue, T. and Rabin, M. (1999a). Doing it now or later. *American Economic Review*, 89(1):103–124.

O'Donoghue, T. and Rabin, M. (1999b). Incentives for procrastinators. *Quarterly Journal of Economics*, 114(3):769–816.

O'Donoghue, T. and Rabin, M. (2001). Choice and procrastination. *Quarterly Journal of Economics*, 116(1):121–160.

Ramsey, F. P. (1928). A mathematical theory of saving. *Economic Journal*, 38:543–559.

Rust, J. (1994). Structural estimation of Markov decision processes. In Engle, R. F. and McFadden, D. L., editors, *Handbook of Econometrics*, volume IV, pages 3082–3139.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 23:165–180.

Wald, A. (1945). Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16(2):117–186.

Weisbrod, B. A. (1964). Collective-consumption services of individual-consumption goods. *The Quarterly Journal of Economics*, 78(3):471–477.

Parametric Family	β	*Mean*	*Std. Deviation*	*Log-Likelihood*
Uniform Naive	1.	-1.86762	5.78115	-1.59186
Uniform Sophisticate	1.	-2.04179	1.87369	-1.59186
Normal Naive	0.82	0.0942045	3.47898	-1.59187
Normal Sophisticate	0.83	0.0978794	3.10058	-1.59187
Extreme Value Naive	0.81	-2.05785	2.37227	-1.59186
Extreme Value Sophisticate	0.83	-1.84762	1.85227	-1.59187
Logistic Naive	0.76	0.193664	9.44528	-1.59187
Logistic Sophisticate	0.77	0.105082	4.10288	-1.59188
Laplace Naive	0.64	0.206991	8.82003	-1.59199
Laplace Sophisticate	0.65	0.0614326	2.24342	-1.59204

Table: Log-likelihood estimates of β and the mean and standard deviation for the example if the analyst does not know the mean and standard deviation of the payoff distribution.

Parametric Family	β	*Log-Likelihood*
Uniform Naive	1.	-3.29153
Uniform Sophisticate	1.	-3.29153
Normal Naive	0.871612	-3.29198
Normal Sophisticate	0.88423	-3.29228
Extreme Value Naive	0.765061	-3.29383
Extreme Value Sophisticate	0.792468	-3.29483
Logistic Naive	0.814908	-3.29203
Logistic Sophisticate	0.836259	-3.29254
Laplace Naive	0.758422	-3.29317
Laplace Sophisticate	0.787311	-3.29418

Table: Log-likelihood estimates of β for the payoff distribution and parameters specified in the example if the analyst knows the mean and standard deviation of the payoff distribution with $T = 30$ periods.

Parametric Family	β	*Log-Likelihood*
Uniform Naive	1.	-3.95505
Uniform Sophisticate	1.	-3.95505
Normal Naive	0.889306	-3.95576
Normal Sophisticate	0.903474	-3.95624
Extreme Value Naive	0.801094	-3.95715
Extreme Value Sophisticate	0.8301	-3.95833
Logistic Naive	0.835118	-3.95584
Logistic Sophisticate	0.85936	-3.9566
Laplace Naive	0.794377	-3.95701
Laplace Sophisticate	0.824827	-3.95823

Table: Log-likelihood estimates of β for the payoff distribution and parameters specified in the example if the analyst knows the mean and standard deviation of the payoff distribution with $T = 60$ periods.

Parametric Family	β	*Log-Likelihood*
Uniform Naive	1.1051	-1.61023
Uniform Sophisticate	1.10823	-1.61029
Normal Naive	1.02514	-1.60953
Normal Sophisticate	1.0253	-1.60953
Extreme Value Naive	1.1942	-1.61034
Extreme Value Sophisticate	1.19231	-1.61008
Logistic Naive	1.	-1.60944
Logistic Sophisticate	1.	-1.60944
Laplace Naive	0.959755	-1.61017
Laplace Sophisticate	0.960106	-1.61016

Table: Log-likelihood estimates of β if the true distribution is Logistic and has the same mean and standard deviation as in the example. We suppose the analyst knows the mean and standard deviation of the payoff distribution, and that $T = 5$ periods.

Parametric Family	β	*Log-Likelihood*
Uniform Naive	0.9	-1.57692
Uniform Sophisticate	0.900684	-1.57692
Normal Naive	0.725994	-1.57692
Normal Sophisticate	0.730595	-1.57693
Extreme Value Naive	0.467228	-1.58092
Extreme Value Sophisticate	0.477292	-1.58106
Logistic Naive	0.670309	-1.57692
Logistic Sophisticate	0.676695	-1.57693
Laplace Naive	0.545986	-1.57699
Laplace Sophisticate	0.555965	-1.57705

Table: Log-likelihood estimates of β for the mean and standard deviation from the example if the agent is naive and $\beta = 0.9$, the true distribution is Uniform, and the analyst knows the mean and standard deviation of the payoff distribution with $T = 5$ periods.

Parametric Family	β	*Mean*	*Std. Deviation*	*Log-L*
Uniform Naive	0.899999	-0.0000121032	3.08835	-1.57
Uniform Sophisticate	0.901039	0.00221368	0.838862	-1.57
Normal Naive	0.729808	0.0281063	2.91605	-1.57
Normal Sophisticate	0.736594	0.0731089	4.76987	-1.57
Extreme Value Naive	0.706168	-0.347689	0.621169	-1.57
Extreme Value Sophisticate	0.633785	0.144273	0.652626	-1.60
Logistic Naive	0.6741	0.0166023	2.176	-1.57
Logistic Sophisticate	0.683439	0.0773394	5.63958	-1.57
Laplace Naive	0.55626	0.017136	1.21714	-1.57
Laplace Sophisticate	0.569426	0.0941048	5.09827	-1.57

Table: Log-likelihood estimates of β, the mean, and standard deviation if the agent is naive and $\beta = 0.9$, the true distribution is Uniform with parameters as in the example, and the analyst does not know the mean and standard deviation of the payoff distribution with $T = 5$ periods.

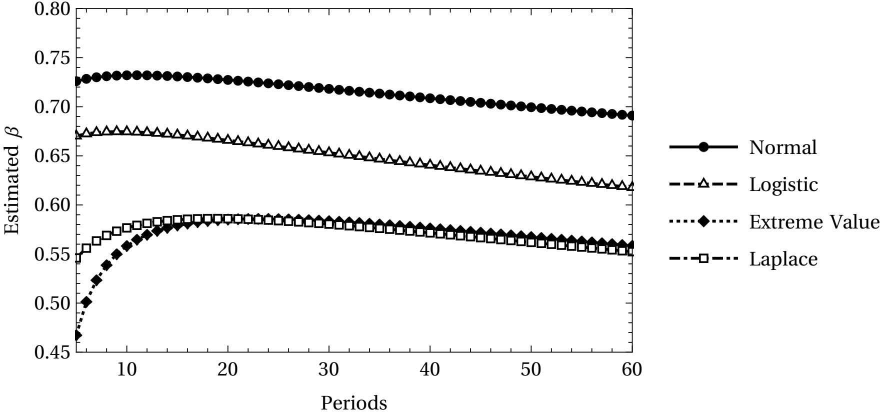


Figure: Estimates of β in the example when the agent is naive and time-inconsistent with $\beta = 0.9, \hat{\beta} = 1, \delta = 1$ for different number of periods T under different parametric assumptions. The analyst knows that $\delta = 1, \hat{\beta} = 1$, as well as the mean and standard deviation of the shock distribution, and estimates β. As the analyst observes the behavior in more and more periods, the estimated value of β eventually moves further away from the true value of 0.9.