

Why don't present-biased agents make commitments?

By DAVID LAIBSON*

*Dept. of Economics, Harvard University, Cambridge, MA 02138 (dlaibson@harvard.edu). The author thanks Nava Ashraf, Daniel Benjamin, Gabriel Carroll, and James Choi for insightful feedback. Outstanding research assistance was provided by Layne Kirshon, Peter Maxted, and Sean Wang. This research was supported by the Pershing Square Fund for Research on the Foundations of Human Behavior and the National Institute on Aging of the National Institutes of Health grant awards R01AG021650 and P01AG005842. The content is solely the responsibility of the author and does not necessarily represent the official views of NIA, NIH, or the Pershing Square Foundation.

Since self-control problems were first analyzed by Strotz (1956), researchers have frequently emphasized that dynamically inconsistent preferences, such as present-biased preferences, engender a demand for commitment.¹ Here, and throughout this paper, I define commitment as a “pure” restriction on one’s choice-set with *no* confounding extrinsic benefits such as tax deferral in a savings plan or intra-household strategic advantages.²

Commitment is a problematic prediction, since we see so little of it in the economy. Researchers have been able to induce some experimental participants to commit themselves (e.g., Ashraf, Karlan, and Yin 2006), but across

a growing literature it is usually the case that only a minority of experimental subjects choose to tie their own hands.³ Moreover, subjects rarely express a willingness to *pay* a significant price to have their choice-set reduced.⁴ Most importantly, very little commitment has arisen in the marketplace without the direct involvement of behavioral economists or their students.⁵

In the current paper, I quantitatively explore the reasons for the “missing” commitment. Extending the present-biased, procrastination model in Carroll et al. (2009), I show how equilibrium commitment is related to (i) the standard deviation of the opportunity cost of time, (ii) the cost of delay, (iii) the degree of partial naiveté, and (iv) the direct cost of commitment.

My quantitative analysis implies that commitment is not a robust implication of present-biased discounting. Once one calibrates the model, commitment vanishes in many leading

¹ This issue plays a central role in Phelps and Pollak (1968), Laibson (1997), and O’Donoghue and Rabin (1999). For related analyses see Gul and Pesendorfer (2001), Fudenberg and Levine (2006), and Amador, Werning, and Angeletos (2006). For a review of the literature on commitment see Bryan, Karlan, and Nelson (2010).

² E.g., Ashraf (2009).

³ For example, see Giné et al. (2010) and Kaur et al. (forthcoming). One study that finds widespread commitment is Beshears et al. (2015).

⁴ For example, see Augenblick et al. (2014), where commitment is popular at a zero price but not at a strictly positive price. One exception is Schilbach (2015).

⁵ stickK was founded by Ian Ayres, Dean Karlan, and a Yale MBA student, Jordan Goldberg, that Ayres and Karlan recruited.

cases. In other words, the benefits of commitment (as perceived by the present-biased agent) are frequently overwhelmed by the costs of commitment.

This does not imply that we should *never* expect to see commitment. Rather, in some natural settings (like the one studied here), commitment is a hothouse flower that survives only under special parameterizations. A demand for commitment is a special case rather than the general case.

Section I explains the basic model (with sophisticated beliefs) and solves it under the assumption that commitment is *not* available. Section II introduces a *free* commitment technology and characterizes the cases under which commitment will be chosen. Section III extends that analysis under the assumption of partial naiveté. Section IV completes the analysis by studying the case in which commitment has a direct cost – i.e., a hassle cost or a market price for setting up a commitment contract. Section V concludes. An associated NBER working paper contains proofs.

I. Model Without Commitment

I extend the model developed in Carroll et al. (2009). The original model has the following features.

Time is discrete, $t \in \{1, 2, 3, \dots\}$. An agent has a present-biased discount function, with

present bias parameter β and $0 < \beta < 1$. The agent has a long-run discount factor $\delta = 1$.

A non-divisible task needs to be done and the agent decides when to do the task. Doing the task requires a single period of effort; if effort is expended during period t , the agent pays an effort cost for that period, c_t . The effort cost for period t , c_t , is realized at the start of period t (so its realized value is not known before period t but is known in period t before the agent decides whether or not to do the task). The effort cost, c_t , is identically and independently distributed each period and it is drawn from a uniform distribution on the interval $[\underline{c}, \bar{c}]$, with CDF $F(c)$. Accordingly, the standard deviation of the distribution is $\sigma = \frac{1}{2\sqrt{3}}[\bar{c} - \underline{c}]$. I assume that the stochastic effort costs, $\{c_1, c_2, c_3, \dots\}$, are non-contractible.

Every period that the task remains undone the agent pays a delay cost (Loss) of $L > 0$. When the agent does the task (in other words, when the agent pays effort cost c_t), the agent stops experiencing any future flow losses.

There exists a Markov equilibrium: act if and only if $c_t \leq c^*$, so c^* is the action threshold. To characterize c^* it is helpful to study the long-run (undiscounted) cost function, $V(c_t)$:

$$V(c_t) = \begin{cases} c_t & \text{if } c_t \leq c^* \\ L + E_t V(c_{t+1}) & \text{if } c_t > c^* \end{cases}$$

Hence, $E_{t-1}V(c_t) =$

$$\int_{c_t \leq c^*} c_t dF(c_t) + \int_{c_t > c^*} [L + E_t V(c_{t+1})] dF(c_t).$$

Because V is a cost function, the agent would like future selves to minimize V , but dynamic inconsistency implies that such cost minimization will not arise in equilibrium.

Instead, in equilibrium, the cutoff rule is characterized by

$$c^* = \beta[L + E_t V(c_{t+1})].$$

Carroll et al. (2009) show that the equilibrium cutoff rule is

$$c^* = \frac{\underline{c} + \sqrt{\underline{c}^2(1 - \beta)^2 + 2\beta L(\bar{c} - \underline{c})(2 - \beta)}}{2 - \beta}.$$

II. Demand for Commitment in the Case of Sophistication

I now turn to an analysis of (self) commitment, which is related to the planner's problem in Carroll et al. (2009). In this section I consider the problem faced by an agent at time "zero," who is deciding whether or not to commit her future selves to binding deadlines. Here I assume that period one is the earliest period that the project can be done, so period zero is a pre-period where the only decision is whether or not to choose deadlines for future selves.

In this section, I consider the case of (perfect) sophistication and a vanishingly small direct price, p , of implementing a commitment contract (i. e., $p \downarrow 0$). An agent will commit to a deadline when the payoff from commitment exceeds the payoff from allowing future selves to have the flexibility to decide when to do the task. For this problem, the personal optimum is either to commit (during period 0) to do the task in period 1, or to allow all future selves to decide for themselves.⁶

I show that commitment will be chosen when

$$\begin{aligned} \frac{(\beta - 1)E[c] + \sigma\sqrt{3}}{2 - \beta} &< L \\ &< \frac{(1 - \beta)E[c]}{\beta} + \frac{\sigma\sqrt{3}}{\beta}. \end{aligned}$$

The first inequality is the threshold at which self zero would like to commit all future selves to act immediately. The second inequality is the threshold at which such commitment becomes redundant because at this threshold the agent will always act immediately for all values of c in the support of F . In other words, the second inequality is the threshold at which $c^* \geq \bar{c}$, so even a vanishingly small direct price of commitment will eliminate commitment.

Figure 1 plots the region described by this pair of inequalities (in the positive orthant), which is labelled "Commitment." Figure 1 also

⁶ See the NBER working paper version of Carroll et al. (2009).

plots the region in which the agent completes the task immediately even if there has not been a commitment (the northwest region labelled “Immediate Action”). Finally, Figure 1 plots the region in which the agent prefers to give her future selves the freedom to decide when to do the task (the southeast region labelled “Procrastination”).

A few properties are apparent in Figure 1. First, as σ increases (holding L fixed and moving horizontally), flexibility/procrastination eventually dominates commitment. Intuitively, the more uncertain the future opportunity cost of time (σ), the less valuable it is to commit to a specific deadline. Second, as L increases (holding σ fixed and moving vertically), commitment eventually ceases to dominate flexibility, because the agent expects to do the activity next period *without* the need for a deadline.

The quantitative values in Figure 1 depend on calibrated values: $\beta = 0.7$ and $E[c] = \$20$. The calibration of present bias is based on typical estimates in the present bias literature.⁷ The calibration of $E[c]$ is based on the joint assumption that the task will take one hour and that the mean opportunity cost of time is a typical hourly wage. Figure 1 also plots the point $(\sigma\sqrt{3} = 20, L = 5)$. This is the case in which

the *standard deviation* of the opportunity cost of an hour of time is $\sigma = \$20/\sqrt{3} = \11.55 . Setting $L = 5$ implies that the household loses \$5 per period for as long as the project remains uncompleted. If periods are days, this amounts to \$1,825 of costs resulting from a year of procrastination on this task.⁸

As you can see, at the point $(\sigma\sqrt{3} = 20, L = 5)$ the agent prefers to procrastinate rather than to commit. But this is only an illustrative example. It is possible to generate reasonable calibrated examples with commitment as the preferred choice – i.e., calibrated points that lie in the shaded “Commitment” region.

III. Demand for Commitment in the Case of Partial Naiveté

Partial naiveté (O’Donoghue and Rabin 2001) weakens the demand for commitment. We can study this weakening quantitatively, using the equations that we have already derived. Specifically, replace β by $\hat{\beta}$ (the agent’s naive expectation of her *future* present bias parameter). Now the band of commitment narrows – see Figure 2 for the case $\hat{\beta} = 0.85 > \beta = 0.7$.

⁷ For example, see Angeletos et al. (2001) and Laibson et al. (2007).

⁸ This is calibrated for a typical household procrastinating on joining a 401(k) plan with a 6% match threshold and a 50% match. For

simplicity, the loss is interpreted to be the lost match, or \$5 per day for a household earning $\frac{365 \times \$5}{0.5 \times 0.06} = \$60,833$.

If $\hat{\beta}$ were raised to one, the two thresholds in Figure 2 would converge to the 45 degree line and the commitment region would vanish.

I don't believe that many economic actors have complete naiveté ($\hat{\beta} = 1$), so I am prone to believe that naiveté is partial (as plotted in Figure 2) and therefore provides only a partial explanation for the lack of equilibrium commitment.

IV. Demand for Commitment When Commitment has a (Non-Zero) Price

Now I study the case in which a commitment contract has a non-trivial implementation price: $p > 0$. This price includes all hassle costs – e.g., taking the time to set up a contract and the system of enforcement – as well as direct payments made to obtain the commitment contract. The costs reflected in p include only the direct cost of setting up the commitment contract (not the indirect cost of lost flexibility). The commitment thresholds are now given by $L =$

$$\frac{\bar{c} \pm \sqrt{\bar{c}^2(1 - \beta)^2 - 2\beta p(\bar{c} - \underline{c})(2 - \beta)}}{\beta(2 - \beta)}$$

$$- E[c] - p.$$

For the calibrated model, the introduction of p turns out to swamp the demand for commitment. Figure 3 illustrates this point, by plotting the commitment region for our calibrated problem ($\beta = 0.7$, $E[c] = \$20$) when the price of the

commitment contract is $p = \$5$. This commitment price reflects a crude estimate of what it would cost a person to set up a commitment contract (including both the internal hassle cost and any revenue paid to a for-profit ‘commitment-services’ firm). The commitment region has now significantly shrunk. It only exists near the origin *and* for very large values of σ and L (that are too large to appear in this figure). For this calibration, a necessary condition for the ‘large- σ ’ solution is $\sigma\sqrt{3} > 159.7$.

Moreover, the collapse of commitment is even more extreme under the assumption of partial naiveté. If we assume that $\hat{\beta} = 0.85$, then the commitment region in the neighborhood of $\sigma = 0$ completely vanishes when the price of the commitment contract is at least \$3.53 and the necessary condition for the ‘large- σ ’ solution is $\sigma\sqrt{3} > 828.4$. Partially naive agents perceive even less reason for commitment than their sophisticated cousins.

Why does the commitment region dramatically shrink at even a modest price, p , for the commitment contract? Commitments are not generating substantial perceived welfare gains. Sophisticates don't gain much from commitment because their welfare losses from procrastination aren't very large. Naifs don't *perceive* that they gain much from commitment, because they don't realize how much their procrastination is (probabilistically) going to hurt them.

Finally, everyone – both sophisticates and naifs – recognize the costs that come with commitment, including the loss of flexibility and the direct price of the commitment contract itself. These costs often swamp the perceived benefits from commitment.

V. Conclusion

These calculations provide a quantitative analysis of the *perceived* benefits of commitment in a particular task completion problem. In this environment, the perceived net benefits from informationally feasible⁹ commitments are modest, implying a weak motive for commitment. In the case that I study, a small price of commitment can tip the scales against commitment. Commitments solve one problem – in this example, procrastination – but produce other problems – time allocations that are insensitive to the opportunity cost of time.

The perceived benefits of commitment will vary across environments. It is an open question if there are many *economically realistic* environments in which informationally feasible commitment contracts have high perceived values.

My quantitative calculations imply that present-biased agents will frequently not make

commitments because the perceived benefits of commitment do not exceed the (modest) direct price of commitment and the indirect losses arising from reduced flexibility. However, the calculations imply that free (or nearly free) commitment technologies may succeed in attracting voluntary adoption.

REFERENCES

- Amador, Manuel, Iván Werning, and George-Marios Angeletos. 2006. “Commitment vs. Flexibility.” *Econometrica* 74 (2): 365-96.
- Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg. 2001. “The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation.” *Journal of Economic Perspectives* 15 (3): 47-68.
- Ashraf, Nava. 2009. “Spousal Control and Intra-Household Decision Making: An Experimental Study in the Philippines.” *American Economic Review* 99 (4): 1245-77.
- Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006. “Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines.” *Quarterly Journal of Economics* 121 (2): 635-72.
- Augenblick, Ned, Muriel Niederle, and Charles

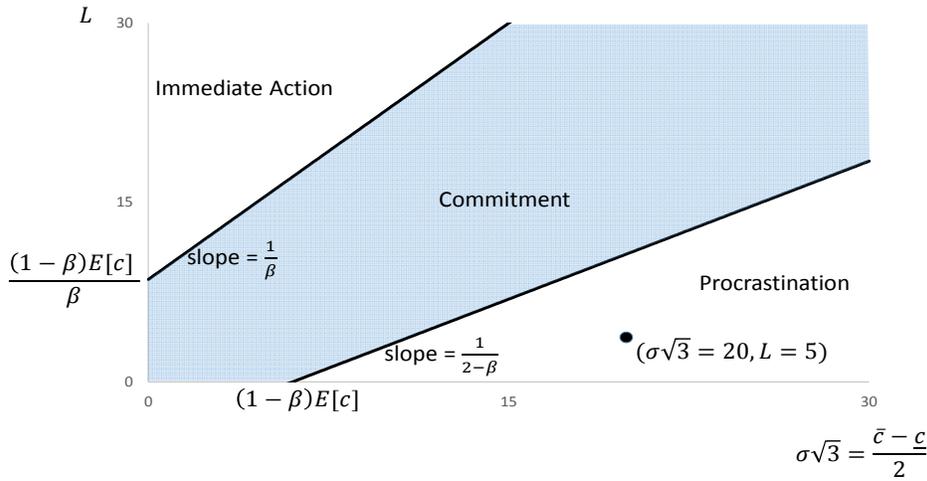
⁹ *Informationally feasible* commitments respect the information asymmetries that exist in the economy – e.g., the fact that a person’s

moment by moment opportunity cost of time is private (non-contractible) information.

- Sprenger. 2014. "Working Over Time: Dynamic Inconsistency in Real Effort Tasks." <http://web.stanford.edu/~niederle/ANS.Paper.pdf>
- Beshears, John, James J. Choi, Christopher Harris, David Laibson, Brigitte C. Madrian, and Jung Sakong. 2013. "Self-Control and Liquidity: How to Design a Commitment Contract." <http://faculty.som.yale.edu/jameschoi/commitment.pdf>.
- Bryan, Gharad, Dean Karlan, and Scott Nelson. 2010. "Commitment Devices." *Annual Review of Economics* 2 (1): 671-98.
- Carroll, Gabriel D., James J. Choi, David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2009. "Optimal Defaults and Active Decisions." *Quarterly Journal of Economics* 124 (4): 1639-74.
- Fudenberg, Drew, and David K. Levine. 2006. "A Dual-Self Model of Impulse Control." *American Economic Review* 96 (5): 1449-76.
- Giné, Xavier, Dean Karlan, and Jonathan Zinman. 2010. "Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation." *American Economic Journal: Applied Economics* 2 (4): 213-35.
- Gul, Faruk, and Wolfgang Pesendorfer. 2001. "Temptation and Self-Control." *Econometrica* 69 (6): 1403-35.
- Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan. Forthcoming. "Self-Control at Work." *Journal of Political Economy*.
- Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics* 112 (2): 443-77.
- Laibson, David, Andrea Repetto, and Jeremy Tobacman. 2007. "Estimating Discount Functions with Consumption Choices over the Lifecycle." NBER Working Paper 13314.
- O'Donoghue, Ted, and Matthew Rabin. 1999. "Doing It Now or Later." *American Economic Review* 89 (1): 103-24.
- O'Donoghue, Ted, and Matthew Rabin. 2001. "Choice and Procrastination." *Quarterly Journal of Economics* 116 (1): 121-60.
- Phelps, E. S., and R. A. Pollak. 1968. "On Second-Best National Saving and Game-Equilibrium Growth." *Review of Economic Studies* 35 (2): 185-99.
- Schilbach, Frank. 2015. "Alcohol and Self-Control: A Field Experiment in India." http://scholar.harvard.edu/files/schilbach/files/jmp_frank_schilbach.pdf.
- Strotz, Robert H. 1956. "Myopia and Inconsistency in Dynamic Utility Maximization." *Review of Economic Studies* 23 (3): 165-80.

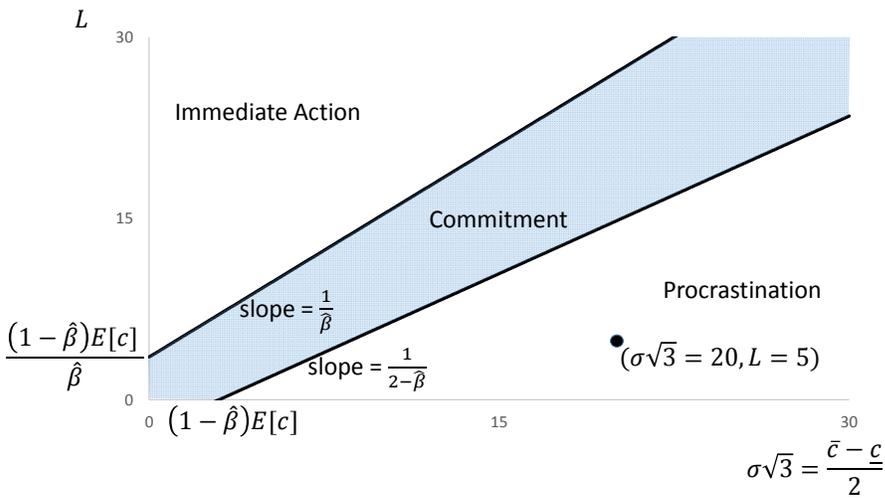
[[Note to illustrator: would you put the three figures on one page, approximating the formatting that I used on the next page. That way they can be viewed comparatively. Is that possible? Or some variant that puts them side by side on one page? Thanks in advance for whatever you can do.]]

FIGURE 1. COMMITMENT REGION FOR A SOPHISTICATED AGENT



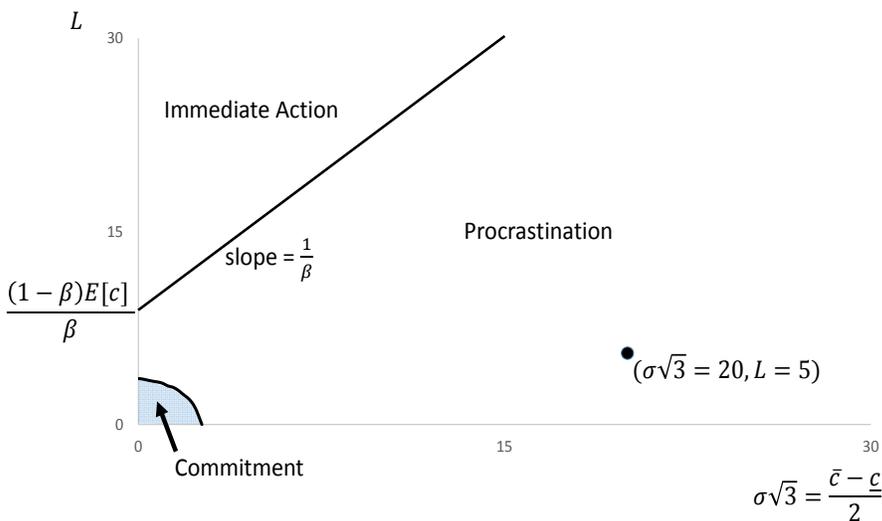
Note: The Immediate Action, Commitment, and Procrastination regions for a sophisticated present-biased agent with $\beta = 0.7$ (and no direct cost for creating/implementing a commitment contract). The horizontal axis is the standard deviation of the opportunity cost of time (scaled by $\sqrt{3}$). The vertical axis is the loss per period from delaying action.

FIGURE 2. COMMITMENT REGION FOR A PARTIALLY NAIVE AGENT



Note: The Immediate Action, Commitment, and Procrastination regions for a partially naive present-biased agent with $\hat{\beta} = 0.85 > \beta = 0.7$ (and no direct cost for creating/implementing a commitment contract). The horizontal axis is the standard deviation of the opportunity cost of time (scaled by $\sqrt{3}$). The vertical axis is the loss per period from delaying action.

FIGURE 3. COMMITMENT REGION FOR A SOPHISTICATED AGENT WITH A DIRECT COST OF COMMITMENT



Note: The Immediate Action, Commitment, and Procrastination regions for a sophisticated present-biased agent with $\beta = 0.7$ and a \$5 direct cost for creating/implementing a commitment contract. The horizontal axis is the standard deviation of the opportunity cost of time (scaled by $\sqrt{3}$). The vertical axis is the loss per period from delaying action.