

## A Appendix

### A.1 Reference-Dependence Anchors

In the most general version of this model  $\lambda$  should be differentiated into  $\lambda_m^b$  ( $\lambda$  specific to losses over bequests) and  $\lambda_m^c$  ( $\lambda$  specific to losses over consumption) to examine how preference and behaviors change when agents perceive a loss to neither, one, or both consumption and bequests. However, for the purposes of this paper, I restrict consumption to always be coded as a gain ( $c \geq 0$ ). This allows me to examine how the transmission of wealth changes only as a result of preferences over bequests, with reference. I still include loss-averse preferences over consumption as it allows the marginal rate of substitution between consumption and bequests to be the same as Cobb-Douglas preferences when agents do not feel a loss over bequests, which is a useful assumption.

In this paper I pin the relative loss-averse anchor as an agent’s own initial wealth. There is an abundance of evidence to support this parameterization of  $\rho_b$ . For example, ? find that parents anchor their expectations of their child’s future education attainment compared to their own. And Barone et al. [2021] find that empirical evidence that social class and loss aversion heavily affects parent’s investments in their children, including parents that are well-off and well-educated. Across the wealth and income spectrum, parents anchor their expectations for their child’s future outcomes based off their own, and will invest more to make sure their children meet that expectation.

There is also supporting evidence that loss-averse anchors are distributionally-dependent for different behaviors and products. For example, Malloy [2015] finds that consumption anchors are dependent on the overall expectation and observation of other’s consumption. So, as the average consumption increases, so too would the reference point. Another good alternative for a distribution-dependent anchor point is on average wealth, an anchoring that teases at the “American dream”. In Ryan et al. [2024], there’s evidence that parent’s anchor their beliefs in what their children should received based on the average wealth of the economy. In times of growth and mobility, this expectation rises; in times of stagnation, the expectation lowers. This would create interesting dynamics that should be explored.

In my model, the assumption that loss averse reference points are only dependent on one’s wealth levels also allows the Markov process to be ergodic, and we can explicitly define the long-run distribution only dependent on the initial distribution. If references are anchored in the distribution, then the process is non-ergodic, and the analysis significantly complicates. This is an important and empirically backed alternative, and further research should explore the effects of non-ergodic processes due to distributional-dependent references in poverty models.

### A.2 Toy Models

#### A.2.1 Linear Production Technology

Let’s assume a linear production technology,  $V(k^*) = k^*R$  as the first-best level of capital and  $V(k) = \underline{k}R$  for  $k < k^*$  as the second-best, and  $R$  is the return rate  $r < R$ . The monitoring function

is piecewise:  $\pi(k) = 0$  if  $k < \bar{k}$  and  $\pi(k) = \pi$  if  $k \geq \bar{k}$ . The incentive-compatibility condition is therefore:  $k(R - r) + wr + T \geq kR - kR\pi$ . Rearranging after plugging in  $k^*$ , the first-best cut-off is  $w^* = k^* - \frac{k^*R\pi + T}{r}$ . Income is therefore piecewise:

$$I(w) = \begin{cases} k^*(R - r) + wr + T & \text{if } w \geq w^* \\ \underline{k}(R - r) + wr + T & \text{if } w < w^* \end{cases}.$$

To calculate the bequest, a poor or rich agents that feels a gain will bequest  $\gamma I(w)$ . If they feel a loss, they increase their bequest to  $\alpha I(w)$ . In the either the poor or rich branches wealth,  $w < w^*$  or  $w \geq w^*$ , we can calculate the sticky ranges of wealth explicitly by formulating the crossings of the 45-degree line at the different shares of income,  $\gamma I(w)$  or  $\alpha I(w)$ . Piecing these together, we get the bequest function in Proposition (??)

**Proposition 4** (Loss-Averse Bequests with Linear Production Function). *Given the assumptions previously defined, we can explicitly define the kinks in the piecewise bequest function:  $w_p = \frac{\gamma}{1-\gamma r} (\underline{k} \cdot (R - r) + T)$  and  $w_1 = \frac{\alpha}{1-\alpha r} (\underline{k} \cdot (R - r) + T)$ . Further,  $w_r = \frac{\gamma}{1-\gamma r} (k^* \cdot (R - r) + T)$  and  $w_3 = \frac{\alpha}{1-\alpha r} (k^* \cdot (R - r) + T)$ . Then, by Definition 2.1,  $\mathcal{S}(w < w^*) = \{[w_p, w_2], [w_r, w_4]\}$ . Thus,*

$$b(w) = \begin{cases} \alpha (k^* \cdot (R - r) + wr + T), & w_3 < w \leq \bar{w} \\ w, & w_r \leq w \leq w_3 \\ \gamma (k^* \cdot (R - r) + wr + T), & w^* \leq w \leq w_r \\ \alpha (\underline{k} \cdot (R - r) + wr + T), & w_1 < w < w^* \\ w, & w_p \leq w \leq w_1 \\ \gamma (\underline{k} \cdot (R - r) + wr + T), & \underline{w} \leq w < w_p \end{cases} \quad (13)$$

and the  $w^* = w_u$ .

*Proof.* To calculate the poor steady state, we calculate when the gain bequest crosses the 45 degree line:  $w = \gamma[\underline{k}(R - r) + wr + T]$ . Rearranging, we obtain  $w_p = \frac{\gamma}{1-\gamma r} (\underline{k} \cdot (R - r) + T)$ . The poor branch bequest at a loss crosses the 45-degree line with propensity  $\alpha$ ,  $w_2 = \frac{\alpha}{1-\alpha r} (\underline{k} \cdot (R - r) + T)$ , and everything between  $w_p$  to  $w_1$  is sticky, and everything above  $w_1$  is a bequest with  $\alpha$  propensity. The same exercise is done for the rich branch.

Since  $0 \leq \underline{k} < k^*$ , there is a discontinuous jump in the income function and therefore the bequest function. Because the change of the income function is constant,  $I'(w) = r$ , then  $\alpha r < 1$  and the income function increases the distance from the 45-degree line indefinitely, so no upper poor sticky interval exists. The jump in income at  $w^*$  is also  $w_u$  if  $k^*$  sufficiently high, which I assume.  $\square$

The function is numerically solved and visualized in Figure 12. The light-gray lines represent a bequest at propensity  $\gamma$ , and the dark-gray lines with propensity  $> \gamma$  up to  $\alpha$ . The loss averse

bequest function follows the light-gray line until  $w_p$  transitioning to the loss averse bequest of the dark-gray line.

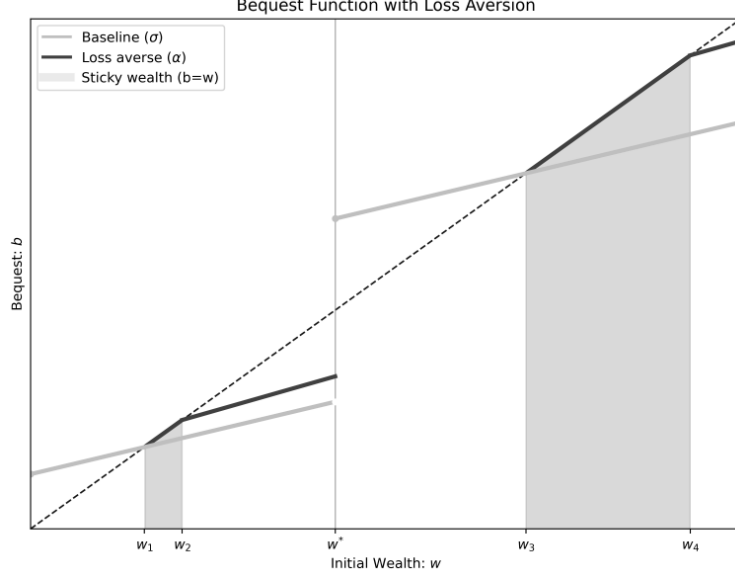


Figure 12: Bequest Function with Loss Aversion (Eq. 13) when  $T > 0$ .

### A.2.2 Convex Production Technology

Given the set-up of Section (2.3.1), we define the incentive comparability condition as:  $k(R - r) + wr + T \geq kR - k^2 R \pi$ . Plugging in  $k^*$ , we rearrange and solve for the first-best borrowers:  $w^* = \underline{k} - \frac{\pi R k^2}{r}$ . For those with  $w < w^*$ , we rearrange the IC as solve for  $k(w)$ , which is

$$k(w) = \frac{r \pm \sqrt{r^2 - 4\pi R(wr + T)}}{2\pi R}.$$

To calculate the sticky regions and to fit the pieces of the bequest function together, we calculate the crossings on the 45-degree line. For gains for the poor, the smaller and larger root of  $k(w)$  represent  $w_p$  and  $w_u$ . For losses for the poor, the smaller and larger root represent the range of the vicious cycle, and in-between the vicious cycle and the gain crossings are sticky intervals. Generally, we calculate this as:

$$\text{Crossings on } [0, w^*]: \quad a_\theta = 4\pi^2 R^2 (\theta r - 1)^2, \quad b_\theta = 8\pi^2 R^2 \theta (\theta r - 1) T + 4\pi R r \theta (R - r) (\theta R - 1),$$

$$c_\theta = 4\theta^2 \pi R T (\pi R T + R(R - r)).$$

$$w_{\theta, \pm} = \frac{-b_\theta \pm \sqrt{b_\theta^2 - 4a_\theta c_\theta}}{2a_\theta}, \quad w_p = \min\{w_{\gamma, -}, w_{\gamma, +}\},$$

$$w_u = \max\{w_{\gamma, -}, w_{\gamma, +}\}, \quad w_1 = \min\{w_{\alpha, -}, w_{\alpha, +}\}, \quad w_2 = \max\{w_{\alpha, -}, w_{\alpha, +}\}.$$

**Crossings on  $[w^*, \infty)$  :**  $w_r = \frac{\gamma [k^*(R-r) + T]}{1 - \gamma r}, \quad w_3 = \frac{\alpha [k^*(R-r) + T]}{1 - \alpha r}.$

We are careful because depending on the parameter, these sticky regions can collapse, expand, or entirely diminish. A sticky region can consume the entire range of  $w_p$  to  $w_u$ , and the general form written above attempts to accommodate the many situations. Assuming nice parameters, we get the result in Proposition (1) and can numerically solve and visualize in Figure (13)

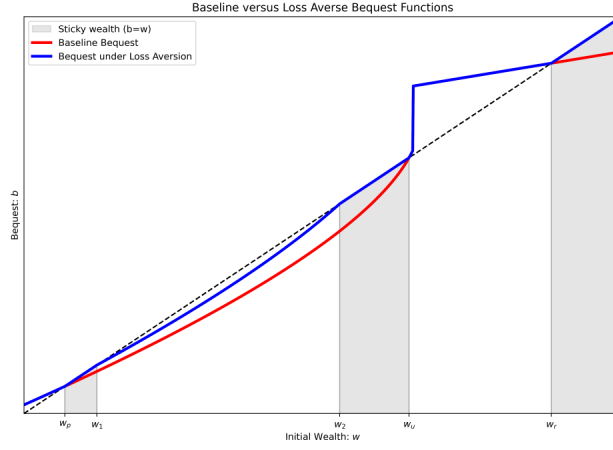


Figure 13: Numerically solved bequest function with convex bequests,  $T > 0$ .

### A.2.3 Collapsing to Canonical Models

**Proposition 5** (Bequest Function with a Linear Production Function). *Let  $\pi(k) = 0$  if  $k < \bar{k}$  and  $\pi(k) = \pi$  for  $k \geq \bar{k}$  where  $0 < \pi \leq 1$ , and  $F > 0$ . Furthermore, I let  $V(k) = Rk$  where  $R$  is the return rate  $r < R$  and  $\bar{k} < k^*$ . Then  $V(k(w)) = 0$  and  $V(k^*) = RL^*$ . Then the baseline bequest function functions takes the form*

$$b_B(w) = \begin{cases} \gamma [k^* \cdot (R-r) + wr + T] & \text{if } w \geq w^* \\ \gamma [wr + T] & \text{if } w < w^* \end{cases} \quad (14)$$

where  $w^* = w_u = k^* - \frac{\pi F + T}{r}$ . The stable steady states are  $w_p = \frac{\gamma T}{1 - \gamma r}$  and  $w_r = \frac{\gamma (k^* (R-r) + T)}{1 - \gamma r}$ .

*Proof.* See appendix. □

In Proposition 5, I assume the most basic structure this model can take and is a useful abstraction. Like in [Banerjee and Newman \[1993\]](#), when the model increases complexity in occupational choice – potentially in risky versus non-risky investment production technologies – the piecewise bequest function can quickly complicate, and a linear technology and monitoring function ensure a tractable model. Since this paper is concerned with the bequest dynamics, the majority of my analysis uses an example of a more sophisticated and realistic production technology and monitoring function outlined in Proposition (6).

**Proposition 6** (Bequest Function with a Convex Production Function). *Let  $V(k) = R \min\{k, \bar{k}\}$  with return rate  $r < R$ . Assume  $\pi(k) = \pi \min\{k, \bar{k}\}$  where  $\pi \in (0, 1)$  and  $\bar{k} < \bar{\bar{k}}$ , and a harsh punishment,  $F = V(k)$ . Then the baseline bequest function takes the form*

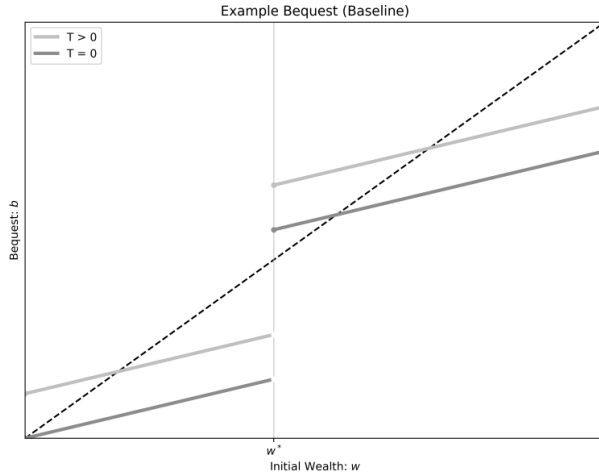
$$b_B(w) = \begin{cases} \gamma[k^* \cdot (R - r) + wr + T] & \text{if } w \geq w^* \\ \gamma[k(w) \cdot (R - r) + wr + T] & \text{if } w < w^* \end{cases} \quad (15)$$

where  $k^* = \bar{k}$ ,  $w^* = k^* - \frac{\pi R(k^*)^2 + T}{r}$  and  $k(w) = \frac{r - \sqrt{r^2 - 4\pi R(wr + T)}}{2\pi R}$ , and  $w_u = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$ . Further,  $w_p = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$  and  $w_r = \frac{\gamma[k^*(R - r) + T]}{1 - \gamma r}$ .

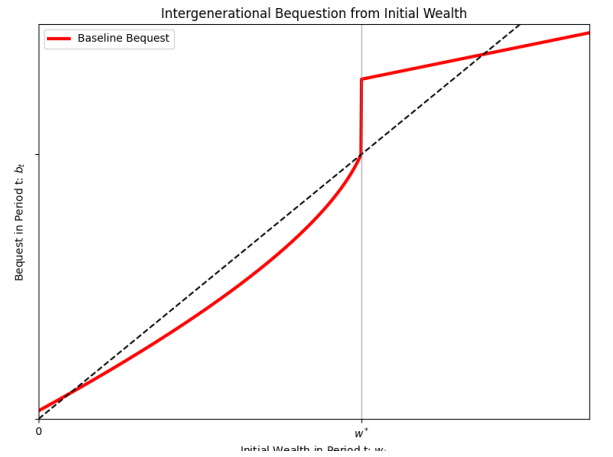
The coefficients are explicitly defined as  $a = 4\pi^2 R^2 (\gamma r - 1)^2$ ,  $b = 8\pi^2 R^2 \gamma (\gamma r - 1) T + 4\pi R r \gamma (R - r) (\gamma R - 1)$ , and  $c = 4\gamma^2 \pi R T (\pi R T + R(R - r))$

*Proof.* See appendix. □

The specifications in Proposition (5) and (6) allow insight into the range of bequest functions possible within the model. The visualizations for these propositions are below. The bequest function in Proposition 5 closely replicates the bequest function in [Banerjee and Newman \[1993\]](#) and the bequest function in Proposition 6 closely replicates the bequest function in [Banerjee and Newman \[1994\]](#). We can visualize the mechanical effects of the market on an agent's bequest in Figure 14b and Figure 14a.



(a) Bequest Function from Propn (5)



(b) Bequest Function from Propn (6)

Figure 14: Example Baseline Bequest Figures

### A.3 Stationary Distribution

$$\begin{aligned} G_\infty = & G_0(w_p)\delta(w_p) + [G(w_1) - G(w_p)] + [G(w_2) - G(w_1)]\delta(w_1) \\ & + [G(w_u) - G(w_2)] + [G(w_r) - G(w_u)]\delta(w_r) + [G(w_3) - G(w_r)] + [1 - G(w_3)]\delta(w_3). \end{aligned}$$

We can simply rearrange this equation to obtain the consolidated stationary distribution.

#### A.4 Empirical Insights, Half-lives

**Half-lives.** The intergenerational wealth transmission speed can be formalized for simulations and potential empirical analysis in Corollary (1). The half-life of a given  $w_0$  is simply number of periods it takes to converge half-way to the corresponding stable steady state, as seen visually in Figure (6).

**Corollary 1** (Bequest Half-Life). *As defined in Proposition (4), for  $w_t < w_2$ , define the half-life to the poor steady state,  $w_p$ , as  $H_i = \frac{\ln(1/2)}{\ln(b'_i(\underline{w}))}$  where  $i$  refers to the baseline model  $B$  or the loss-averse model  $L$ . Then,*

$$H_L = \frac{-\ln(2)}{\ln(\frac{\alpha}{\gamma}) + \ln(b'_B(w_t))} > \frac{-\ln(2)}{\ln(b'_B(w_t))} = H_B.$$

$$H_B = H_L \left( \frac{\ln(2)}{\ln(2) + \ln(\frac{\alpha}{\gamma})H_L} \right)$$

when  $0 < \frac{\alpha}{\gamma} \frac{\ln(2)}{H_L} < 1$ , which is true for  $w \in [w_1, w_2]$ . Otherwise, the  $H_L \rightarrow \infty$  without some idiosyncratic shock.

Using the Corollary (1), if a researcher can estimate the observed half-life in panel data, this model allows them to then calculate the underlying true half-life. This would require estimating two additional components: the propensity to bequest when not feeling a loss and the propensity at a loss. Since this is hard to capture in wealth data, consumption data would allow insight into those different bequest propensities, which is well-analyzed [Fisher et al., 2020]. It is beyond the scope of this paper to do this, but could be a fruitful avenue for research

#### A.5 T1 Robustness Check

I employ a robustness check of the identified thresholds in the loss-averse and baseline models using bootstrapping, which identifies the  $\hat{c}$  that maximizes the F-stat for a given guess of  $c$  in Table 3. While this shows confidence in the 99.75 percentile, this is a somewhat trivial exercise in a Monte Carlo simulation since the data generating function is simple with clear i.i.d stochastic errors. Note that the sup-F for loss aversion, while still high, is about 1/4 the value of the model without loss aversion. As data becomes noisier and less rich, the F-stat for the best  $\hat{c}$  will decrease, and the loss averse model might produce a low F-stat even if there is a poverty trap.

Table 3: T1 threshold existence (sup-F across  $c$ ; bootstrapping  $p$ )

| Model       | $\hat{c}$ (sup-F) | sup-F    | $p$ -value |
|-------------|-------------------|----------|------------|
| Loss-averse | 0.8418            | 107464.9 | 0.0025     |
| Baseline    | 0.8477            | 408392.8 | 0.0025     |

Due to the sticky ranges of wealth and the slower convergence towards steady states, the loss averse model shows that even the reliable threshold test can fail to reject the null hypothesis when it otherwise should in the baseline.

## A.6 Proofs for Optimal Tax in single-period flat tax rebate policy

### Model Set-up

To begin, note that each individual will maximize utility to their updated budget constraint, where *net capital gains* is taxed at rate  $\tau$ , and they receive some lump-sum rebate,  $T$ :

$$\max_{b^i} U(c^i, b^i) = (1 - \gamma) \ln(c^i) + \gamma \ln(b^i) + \eta \cdot \gamma \nu(\ln(b) | \ln(w)) + \eta \cdot (1 - \gamma) \nu(\ln(c) | \ln(0^+)) \quad (16)$$

$$\begin{aligned} \text{s.t. } c^i + b^i &= (1 - \tau)k(w^i, \tau)(R - r) + w^i \cdot r + T \\ c^i, b^i &\geq 0 \end{aligned} \quad (17)$$

This creates a trade-off for borrowers that affects the lenders' new incentive-compatibility constraint:

$$(1 - \tau)k(R - r) + w \cdot r + T = [1 - \pi k] \cdot Rk \quad (18)$$

Rearranging Equation (18), lenders will only administer a small enough such that agents are indifferent to repayment and renegeing, where they choose repayment in the static equilibrium. To borrow the first-best level of capital  $k^*$ , an agent must have  $w^* = \frac{k^*(\tau(R-r)+r)-(k^*)^2(\pi R)-T}{r}$ . Intuitively, as the tax rate goes up, so does the required wealth for first-best borrowers  $k(\tau(R - r) + r) > kr$  but goes down as the rebate goes up  $T > 0$ . Intuitively, as the tax goes up the rebate increases the lowest poverty stable steady state, but the rebate might be small compared to the loss in returns for higher wealth individuals  $T - k^*R\tau < 0$ . For non-first best borrowers, the lending function for those with  $w < w^*$  is

$$k(w, \tau, T) = \frac{\tau(R - r)}{2\pi R} + \frac{r - \sqrt{(\tau(R - r) + r)^2 - 4\pi R(wr + T)}}{2\pi R} \quad (19)$$

Mathematically, while  $T > 0$  is an upward shift in the return function, the tax  $\tau$  causes a clockwise rotation around the y-intercept. Intuitively, while it increases the wealth of agents, the loss of returns also hurt agents borrowing the second-best level of capital. The new bequest function incorporates these trade offs. Solving the MRS given the Utility Function (16) and the income function in Equation (19), the bequest function when agents feel a gain is:

$$I(w, \tau) = \begin{cases} k^* \cdot (1 - \tau)(R - r) + w \cdot r + T, & \text{if } w \geq w^* \\ k(w, \tau) \cdot (1 - \tau)(R - r) + w \cdot r + T, & \text{if } w < w^* \end{cases} \quad (20)$$

By combining the above income in Equation (20) with the utility function in Equation (16), we can rederive the updated bequest function. Instead of rewriting the long bequest function here, note that the only change from Proposition (1) is that net-income is taxed, and the good-behavior  $T$

benefit is the demogrant.

To calculate the demogrant, I first describe the total capital in the economy as  $K$ , which is simply the mass of those borrowing first-best and then the expected value of those earning their individualized second-best level of capital,  $K = \int_{w^*}^{\bar{w}} k^* g(w) \partial w + \int_{\underline{w}}^{w^*} k(w, \tau) g(w) \partial w$ . This can be reduced and utilized in Equation (21):

$$K = k^* \cdot [1 - G(w^*)] + \mathbb{E}[k(w, \tau) \mid w \in [\underline{w}, w^*]] \quad (21)$$

Given aggregate capital in Equation (21), the rebate each agent receives in  $T = \tau K(R - r)$ , since  $G(w)$  is normalized to one.

### Solve the Model (without loss aversion).

The general social planner problem is below. However, I am first going to solve for the case when  $\eta = 0$ , individuals don't feel loss averse. This will allow me to then easily flesh out the SWF for loss averse preferences.

$$\begin{aligned} \mathbb{W} &= \int_i g(w^i) \cdot U^i(w^i, \tau) \partial w^i = \sum_{(a,b)} \int_a^b g(w^i) \cdot U_j^i(w^i, \tau) \partial w^i \\ (a, b) &= \{[\underline{w}, w_p], [w_p, w_1], [w_1, w_2], [w_2, w_u], \dots\} \\ \text{s.t. } \tau K(R - r) &\geq 0 \end{aligned} \quad (22)$$

A social planner only considering the cobb-douglas utility for non-loss averse agents will solve the following equation for  $\tau^*$ :

$$\frac{\partial SWF}{\partial \tau} = \frac{\partial}{\partial \tau} \left( \int_{w^*}^{\bar{w}} U(w^i, \tau) g(w^i) \partial w^i + \int_{\underline{w}}^{w^*} U(w^i, \tau) g(w^i) \partial w^i \right) = 0.$$

To solve this, we begin by solving for the partial derivative of utility. Since agents optimize their utility over bequests,  $\frac{\partial U(b^*)}{\partial b} = 0$ , then using the envelope theorem as described, we see

$$\begin{aligned} \frac{\partial U}{\partial \tau} &= \frac{\partial(\gamma \ln(b^*))}{\partial b} \cdot \frac{\partial b^*}{\partial \tau} + \frac{\partial((1 - \gamma) \ln(I - b^*))}{\partial b} \cdot \left( \frac{\partial I}{\partial \tau} - \frac{\partial b^*}{\partial \tau} \right) \\ &= \left( \frac{\partial(\gamma \ln(b^*))}{\partial b} - \frac{\partial((1 - \gamma) \ln(I - b^*))}{\partial b} \right) \cdot \frac{\partial b^*}{\partial \tau} + \frac{\partial((1 - \gamma) \ln(I - b^*))}{\partial b} \cdot \frac{\partial I}{\partial \tau} \\ &= 0 \cdot \frac{\partial b^*}{\partial \tau} + \frac{1 - \gamma}{I - b^*} \cdot \frac{\partial I}{\partial \tau} \\ &= \frac{1 - \gamma}{(1 - \gamma)I(w, \tau)} \cdot \frac{\partial I}{\partial \tau} \\ U_\tau &= \frac{1}{I(w, \tau)} \frac{\partial I}{\partial \tau} \end{aligned}$$

Let  $L_\tau(w, \tau)$  denote  $\partial k(w, \tau, T)/\partial \tau$  as implied by (19), and  $K_\tau = \int_{\underline{w}}^{w^*} L_\tau(w, \tau) g(w_i) \partial w_i$ . This

allows us to take the derivative of each part of the income function in equation (20):

$$\begin{aligned} I_\tau(w < w^*) &= (R - r) \cdot ((1 - \tau)L_\tau(w, \tau) - k(w, \tau)) + (R - r) \cdot (K + \tau K_\tau) \\ &= (R - r) \cdot ((1 - \tau)L_\tau(w, \tau) - k(w, \tau)) + (R - r) \cdot (K - \tau \frac{\partial K}{\partial(1 - \tau)}) \end{aligned}$$

$$\begin{aligned} I_\tau(w \geq w^*) &= (R - r) \cdot (-k^*) + (R - r) \cdot (K + \tau K_\tau) \\ &= (R - r) \cdot (-k^*) + (R - r) \cdot (K - \tau \frac{\partial K}{\partial(1 - \tau)}) \end{aligned}$$

Intuitively, the first part of the income derivative simply says a marginal increase in the tax will increase the total tax revenue which has a positive influence on income, but decrease the total capital gain. The second parts thus describe that as the marginal tax increases, so too do the intensive marginal effects of income. For small  $k(w, \tau)$ , an increase in the tax has a positive effect with diminishing marginal returns. Eventually this switches. Both describe the mechanical and reaction effects of the tax on income.

Now, we can take the derivative of the welfare function. Given the Leibniz integration rule, we can simply move the  $\tau$  derivative into the integrals. We do not need to differentiate  $\frac{\partial w^*}{\partial \tau}$  since the

measure of agents at that wealth is zero. However, it can easily be derived for an exercise.

$$\begin{aligned}
\frac{\partial W}{\partial \tau} &= \int_{w^*}^{\bar{w}} U_\tau(w^i, \tau) \cdot g(w^i) \partial w^i + \int_{\underline{w}}^{w^*} U_\tau(w^i, \tau) \cdot g(w^i) \partial w^i = 0 \\
0 &= \int_{w^*}^{\bar{w}} (R - r) \frac{-k^* + [K - \tau \frac{\partial K}{\partial (1-\tau)}]}{I(w^i, \tau)} \cdot g(w^i) \partial w^i \\
&\quad + \int_{\underline{w}}^{w^*} (R - r) \frac{(1 - \tau) L_\tau(w, \tau) - k(w, \tau) + [K - \tau \frac{\partial K}{\partial (1-\tau)}]}{I(w^i, \tau)} \cdot g(w^i) \partial w^i \\
0 &= \int_{w^*}^{\bar{w}} \frac{-k^*}{I(w^i, \tau)} \cdot g(w^i) \partial w^i + \int_{\underline{w}}^{w^*} \frac{(1 - \tau) L_\tau(w, \tau) - k(w, \tau)}{I(w^i, \tau)} \cdot g(w^i) \partial w^i \\
&\quad + \int_{w^*}^{\bar{w}} \frac{K - \tau \frac{\partial K}{\partial (1-\tau)}}{I(w^i, \tau)} \cdot g(w^i) \partial w^i + \int_{\underline{w}}^{w^*} \frac{K - \tau \frac{\partial K}{\partial (1-\tau)}}{I(w^i, \tau)} \cdot g(w^i) \partial w^i \\
0 &= \int_{w^*}^{\bar{w}} \frac{-k^*}{I(w^i, \tau)} \cdot g(w^i) \partial w^i + \int_{\underline{w}}^{w^*} \frac{(1 - \tau) L_\tau(w, \tau) - k(w, \tau)}{I(w^i, \tau)} \cdot g(w^i) \partial w^i \\
&\quad + \left[ K - \tau \frac{\partial K}{\partial (1-\tau)} \right] \left( \int_{w^*}^{\bar{w}} \frac{1}{I(w^i, \tau)} \cdot g(w^i) \partial w^i + \int_{\underline{w}}^{w^*} \frac{1}{I(w^i, \tau)} \cdot g(w^i) \partial w^i \right) \\
\left[ K - \tau \frac{\partial K}{\partial (1-\tau)} \right] &\left( \int_{w^*}^{\bar{w}} \frac{1}{I(w^i, \tau)} \cdot g(w^i) \partial w^i + \int_{\underline{w}}^{w^*} \frac{1}{I(w^i, \tau)} \cdot g(w^i) \partial w^i \right) \\
&= \int_{w^*}^{\bar{w}} \frac{k^*}{I(w^i, \tau)} \cdot g(w^i) \partial w^i + \int_{\underline{w}}^{w^*} \frac{k(w, \tau) - (1 - \tau) L_\tau(w, \tau)}{I(w^i, \tau)} \cdot g(w^i) \partial w^i.
\end{aligned}$$

We can define  $e = \frac{1-\tau}{K} \frac{\partial K}{\partial (1-\tau)}$ . Then, we can simplify the integrals in expectation. With a specification of  $G(W)$ , we can explicitly solve for  $\tau^*$ . But with this general form, I offer the implicit solution.

$$\begin{aligned}
K \left[ 1 - \frac{\tau}{1 - \tau} e \right] &\left( \mathbb{E} \left[ \frac{1}{I(w^i, \tau)} \mid w^i \in [w^*, \bar{w}] \right] + \mathbb{E} \left[ \frac{1}{I(w^i, \tau)} \mid w^i \in [\underline{w}, w^*] \right] \right) \\
&= \int_{w^*}^{\bar{w}} \frac{k^*}{I(w^i, \tau)} \cdot g(w^i) \partial w^i + \int_{\underline{w}}^{w^*} \frac{k(w, \tau) - (1 - \tau) L_\tau(w, \tau)}{I(w^i, \tau)} \cdot g(w^i) \partial w^i \\
K \left[ 1 - \frac{\tau}{1 - \tau} e \right] &= \frac{\mathbb{E} \left[ \frac{k^*}{I(w^i, \tau)} \mid w^i \in [w^*, \bar{w}] \right] + \mathbb{E} \left[ \frac{k(w, \tau) - (1 - \tau) L_\tau(w, \tau)}{I(w^i, \tau)} \mid w^i \in [\underline{w}, w^*] \right]}{\mathbb{E} \left[ \frac{1}{I(w^i, \tau)} \mid w^i \in [w^*, \bar{w}] \right] + \mathbb{E} \left[ \frac{1}{I(w^i, \tau)} \mid w^i \in [\underline{w}, w^*] \right]}.
\end{aligned}$$

Let  $G$  be defined as the right-side of the equation, and  $\bar{G} = \frac{G}{K}$

$$K \left[ 1 - \frac{\tau}{1 - \tau} e \right] = G$$

Then we can implicitly solve for the welfare-maximizing tax rate,

$$\tau^* = \frac{1 - \bar{G}}{1 - \bar{G} + e}.$$

This is the same implicit  $\tau^*$ .

### **Solve the Model (with loss aversion).**

Now I allow for loss aversion using the same set up as Equation (22). Much like the additional complications loss aversion creates to the bequest function as seen in Proposition (1), the solution the social welfare problem will be similar the model without loss aversion, only this time there are several more integrals branches instead of just those above and below  $w^*$ . These integrals follow the different branches in the loss averse bequest function. As outlined in Equation (22), we just sum the integrals where those roots exist.

To begin, using the envelop theorem, we can determine that for agents feeling a *gain*, their utility derivative with respect to the tax rate is

$$\begin{aligned} \frac{\partial U(b^* \geq w)}{\partial \tau} &= \frac{1 - \gamma}{(1 - \gamma)I(w, \tau)} \frac{\partial I(w, \tau)}{\partial \tau} + \frac{\eta(1 - \gamma)}{(1 - \gamma)I(w, \tau)} \frac{\partial I(w, \tau)}{\partial \tau} \\ &= \frac{1 + \eta}{I(w, \tau)} \frac{\partial I(w, \tau)}{\partial \tau} \end{aligned}$$

The term  $(1 + \eta)$  increases the welfare, but if all agents were to feel a gain, this term would cancel out of the equation  $\frac{\partial SWF}{\partial \tau} = 0$ , and we return to the world without loss aversion at all. Additionally, if we didn't cancel the term, this would not change the maximizing tax rate, just the value of  $SWF(\tau^*)$  though this is not in-and-of itself comparable to other welfare results.

For agents experiencing a loss, their utility derivative is

$$\begin{aligned} \frac{\partial U(b^* < w)}{\partial \tau} &= \frac{1 - \gamma}{(1 - \alpha)I(w, \tau)} \frac{\partial I(w, \tau)}{\partial \tau} + \frac{\eta(1 - \gamma)}{(1 - \alpha)I(w, \tau)} \frac{\partial I(w, \tau)}{\partial \tau} \\ &= \frac{(1 + \eta)(1 - \gamma)}{(1 - \alpha)I(w, \tau)} \frac{\partial I(w, \tau)}{\partial \tau} \end{aligned}$$

As  $\lambda$  or  $\eta$  increase, so does  $U_\tau(b^*, \tau)$  since  $\frac{1 - \gamma}{1 - \alpha} > 1$ . For  $\lambda > 1$  and  $\eta > 0$ , then marginal effect of a tax rate will be greater under a loss than without. For agents with  $w \geq w^*$ , this means a increase in a tax has a greater *negative* impact. For agents with  $w < w^*$ , this means an increase in a tax has a greater *positive* impact. Thus, the optimal tax becomes more sensitive to the distribution. If

we add one more person to the distribution with initial wealth  $w < w^*$  and who feels a loss, then the tax rate will increase compared to the world without loss aversion; conversely, if we add one more person with  $w \geq w^*$  who feels a loss, the tax decreases

Since income is not dependent on loss averse preferences, the derivative doesn't change. However, since welfare function must account for the different piece of possible bequest branches, which are also utility branches. I will write out the general derivative below, then provide the final expected value since it follow a similar albeit more droning process as the model without loss aversion. There are two camps of individuals that *do not* feel a loss: those poorer than  $w_p$  converging up and those in the range of mobility,  $[w_u, w_r]$ . Those that do feel a loss are those in the trap  $(w_p, w_u)$  and those above  $w_r$ , from  $(w_r, \bar{w}]$ . In each camp, there can be first-best and second-best borrowers. Due to the envelope theorem, the bequest amount is not necessary to consider here when the planner maximizes, so we aren't concerned with the sticky regions.

$$\begin{aligned} \frac{\partial W}{\partial \tau} = & \int_{w_r}^{\bar{w}} U_{\tau}^L(w^i, \tau) \cdot g(w^i) \partial w^i + \int_{w^*}^{w_r} U_{\tau}^G(w^i, \tau) \cdot g(w^i) \partial w^i \\ & + \int_{w_u}^{w^*} U_{\tau}^G(w^i, \tau) \cdot g(w^i) \partial w^i + \int_{\underline{w}}^{w_u} U_{\tau}^L(w^i, \tau) \cdot g(w^i) \partial w^i = 0 \end{aligned}$$

Following this, we can

$$\begin{aligned} & \left[ K - \tau \frac{\partial K}{\partial (1 - \tau)} \right] \cdot \left[ \frac{1 - \gamma}{1 - \alpha} \cdot \mathbb{E} \left[ \frac{1}{I^*(w^i, \tau)} \mid w^i \in (w^*, \bar{w}) \right] + \mathbb{E} \left[ \frac{1}{I^*(w^i, \tau)} \mid w^i \in (w_r, \bar{w}) \right] \right. \\ & \quad \left. \mathbb{E} \left[ \frac{1}{I(w^i, \tau)} \mid w^i \in (w^*, \bar{w}) \right] + \frac{1 - \gamma}{1 - \alpha} \cdot \mathbb{E} \left[ \frac{1}{I(w^i, \tau)} \mid w^i \in (w_r, \bar{w}) \right] \right] \\ & = \left[ \frac{1 - \gamma}{1 - \alpha} \cdot \mathbb{E} \left[ \frac{k^*}{I^*(w^i, \tau)} \mid w^i \in (w^*, \bar{w}) \right] + \mathbb{E} \left[ \frac{k^*}{I^*(w^i, \tau)} \mid w^i \in (w_r, \bar{w}) \right] \right. \\ & \quad \left. \mathbb{E} \left[ \frac{k(w^i, \tau) - (1 - \tau)L_{\tau}(w^i, \tau)}{I(w^i, \tau)} \mid w^i \in (w^*, \bar{w}) \right] + \frac{1 - \gamma}{1 - \alpha} \cdot \mathbb{E} \left[ \frac{k(w^i, \tau) - (1 - \tau)L_{\tau}(w^i, \tau)}{I(w^i, \tau)} \mid w^i \in (w_r, \bar{w}) \right] \right] \end{aligned}$$

While the equation looks very messy and complicated, ultimately this is just a sum of expected values for across those making first-best and second-best, and then those that feel a gain and those feel a loss, so 4 integrals and expected values. For those feeling a gain, regardless of income level, they are waited as in the baseline. For those feeling a loss, they are upweighted.

$$K - \tau \frac{\partial K}{\partial (1 - \tau)} = \frac{\sum_j \mathbb{E}_j [I_{\tau}(w^i, \tau) \mid b^* \geq w^i] + \frac{1 - \gamma}{1 - \alpha} \mathbb{E}_j [I_{\tau}(w^i, \tau) \mid b^* < w^i]}{\sum_j \mathbb{E}_j \left[ \frac{1}{I(w^i, \tau)} \mid b^* \geq w^i \right] + \frac{1 - \gamma}{1 - \alpha} \mathbb{E}_j \left[ \frac{1}{I(w^i, \tau)} \mid b^* < w^i \right]}$$

where  $j$  indicates a sum over the expected values when  $w < w^*$  and  $w \geq w^*$ . In the LHS, the denominator sums the expected values by their loss/gain weight of the recipricol of their branch's income. The numerator sums the expected values of the change of income given a marginal increase

in the tax. Whether the numerator or denominator is larger or smaller is entirely dependent on the distribution and the rebate/tax. For wealth individuals, a marginal increase has a negative effect on the tax, and quickly shrinks the numerator faster than the denominator. For poorer individuals, an increase in tax reduces their income, but the rebate *might* be a stronger benefit. Let  $G$  equal the RHS.

$$\tau^* = \frac{1 - G_L}{1 - G_L + e}. \quad (23)$$

We can visualize the social planner's trade off for the curren