

## Volume 43, Issue 3

### Optimal tax design with costly tax evasion

Siqi Wang

*Graduate School of Economics and Business, Hokkaido  
University*

Jun-ichi Itaya

*Department of Law and Economics, Hokusei Gakuen  
University*

#### Abstract

This paper extends the Atkinson-Stiglitz analysis to a dynamic overlapping generations model, incorporating the realistic assumption that agents can evade labor income taxes by misreporting their true income. They can do so by incurring both non-monetary and monetary costs, which are distinguished in the paper and shown to have different implications for optimal tax policies. By considering the monetary cost as a deferred payment, the paper shows that tax evasion concerns can render the well-known Atkinson-Stiglitz theorem invalid and provides critical insights into how different types of costs affect tax schedules.

---

**Citation:** Siqi Wang and Jun-ichi Itaya, (2023) "Optimal tax design with costly tax evasion", *Economics Bulletin*, Volume 43, Issue 3, pages 1271-1278

**Contact:** Siqi Wang - [siqi.wang.f0@elms.hokudai.ac.jp](mailto:siqi.wang.f0@elms.hokudai.ac.jp), Jun-ichi Itaya - [j-itaya@hokusei.ac.jp](mailto:j-itaya@hokusei.ac.jp).

**Submitted:** April 02, 2023. **Published:** September 30, 2023.

## 1. Introduction

The optimal taxation theory, originally proposed by Mirrlees (1971), has significantly influenced taxation theory by providing a framework that balances efficiency and redistribution. The crux of this theory is that the policy planner can only observe agents' incomes, but not their true earning abilities, leading to discussions about problems with asymmetric information. Casamatta (2021) expands the scope of the asymmetric information problem by explaining "earning income" as "reported income". In such a scenario, the asymmetric information problem is not only due to individuals misreporting their type but also due to the concealment of income, which leads to tax avoidance and evasion.

Recently, several contributions analyze the optimal non-linear taxation problems in non-compliance frameworks. Blomquist, Christiansen, and Micheletto (2016) examined diverse misreporting behaviors and variations in labor supply among heterogeneous agents. They shed light on how the government should address the issue of public provision when agents are capable of engaging in misreporting activities. Gahvari and Micheletto (2020) develop the optimal taxation framework with tax avoidance to a general equilibrium, and they argue that there is an additional tax on the low-skilled and a subsidy on the high-skilled due to the wage endogeneity assumption.

When examining capital taxation within the Mirrlees (1971) framework, the work of Atkinson and Stiglitz (1976) immediately comes to mind. This study concludes that indirect taxation should not be used when there is a separability in preferences between consumption and labor. Applying the Atkinson-Stiglitz theorem to the context of treating consumptions in different periods as distinct commodities leads to the corollary that there should be zero capital tax. To examine optimal capital taxation in the context of non-compliance, Christiansen and Tuomala (2008) introduced a dual tax system that incorporates both non-linear labor income taxation and capital taxation, while allowing agents to shift their incomes. They observed that in a progressive tax system, income shifting from labor to capital could arise, particularly when it is difficult to distinguish between the two types of income.

This paper extends the Atkinson-Stiglitz analysis to a dynamic overlapping generations model by introducing a more realistic assumption that agents can evade labor income taxes by misreporting their true income, incurring both non-monetary and monetary costs. The nonmonetary cost is associated with the disutility of hiding income, such as the effort required or the guilt attached to the behavior, while the monetary cost involves the penalties that may be imposed if agents are caught in the future. To simplify and maintain generality, we adopt

the same assumption of riskless utility as Blomquist, Christiansen, and Micheletto (2016) and consider the monetary cost (penalty) as a tax debt that individuals owe to the government and are obligated to pay in the future. Based on our analysis under these assumptions, it appears that the Atkinson-Stiglitz theorem may no longer hold.

The remainder of this paper is organized as follows: Section 2 introduces the fundamental models. Section 3 delves into the subject of optimal taxation. Finally, Section 4 provides concluding remarks.

## 2. The model

By employing the two-type model introduced by Stern (1982) and Stiglitz (1982), we assume the economy consists of two types of agents in each period distinguished only by their productivities denoted by  $w_t^H$  and  $w_t^L$ , respectively. For a type  $i$  ( $i = H, L$ ) agent born at period  $t$ , his preference is represented by the utility function  $U(c_t^i, x_t^i, l_t^i)$ , where  $c_t^i$  denotes his first-period consumption,  $x_t^i$  expresses the second-period consumptions, and  $l_t^i$  is his labor hours. The agent's utility function is given by:

$$U_t^i = c_t^i - \phi(\Delta_t^i) + v(x_t^i) - \varphi(l_t^i), \quad (1)$$

where  $v(\cdot)$  is strictly increasing and concave, while  $\varphi(\cdot)$  is strictly increasing and convex. Denote reported income by  $M_t^i$  and hidden income by  $\Delta_t^i$ , the agent's labor hours can be rewritten as  $l_t^i = (M_t^i + \Delta_t^i)/w_t^H$ . The non-monetary cost, represented by  $\phi(\Delta_t^i)$ , must be paid instantly when an individual conceals their earned income, while the monetary cost, denoted by  $P(\Delta_t^i)$ , refers to a cost that individuals will face in the future. No cost should be paid if no concealment action happens, i.e.  $\phi(0) = P(0) = 0$ . In addition, we assume both  $\phi(\cdot)$  and  $P(\cdot)$  are continuously increasing on  $[0, +\infty]$ , and convex. Then the budget constraints are given by:

$$c_t^i = M_t^i - T(M_t^i) + \Delta_t^i - (1 + \tau_t)k_{t+1}^i, \quad (2)$$

$$x_t^i = r_{t+1}k_{t+1}^i - P(\Delta_t^i), \quad (3)$$

where  $k_{t+1}^i$  is the capital investment for consumption in the second period while  $r_{t+1}$  is the gross rate of return per capital input.  $T(M_t^i)$  represents a non-linear labor income tax based on the reported income, and  $\tau_t$  represents the linear capital tax rate.

To solve the agents' problems, it is convenient to make use of the two-stage approach as Christiansen (1984). First, we hold the agent's net reported income denoted by  $B_t^i \equiv M_t^i - T(M_t^i)$  constant, which also implies the reported income  $M_t^i$  is fixed. Assuming there are interior solutions and then the first-order conditions on the second stage are given by:

$$1 - \phi'(\Delta_t^i) - v'(x_t^i)P'(\Delta_t^i) = \phi'(l_t^i)/w_t^i, \quad (4)$$

$$(1 + \tau_t) = v'(x_t^i)r_{t+1}. \quad (5)$$

After that, we immediately obtain the best choices bundle  $(\Delta_t^i, k_{t+1}^i)$  and then  $(c_t^i, x_t^i, l_t^i)$ , which are all functions of  $B_t^i$ ,  $M_t^i$ , and  $\tau_t$ , and define the conditional indirect utility function  $V(B_t^i, M_t^i, \tau_t; w_t^i)$ .

In the first stage, labor hours are chosen to maximize the conditional indirect utility function. While labor is denoted by  $l_t^i = (M_t^i + \Delta_t^i)/w_t^i$ , we regard the only choice variable as  $M_t^i$ . The first-order condition in the first stage is then given by:

$$\frac{dV(B_t^i(M_t^i), M_t^i, \tau_t)}{dM_t^i} = V_B^{i,t} \frac{dB_t^i}{dM_t^i} + V_M^{i,t} \equiv V_B^{i,t} (1 - T'(M_t^i)) + V_M^{i,t} = 0. \quad (6)$$

## 2.1 The agents' behavior response

Using the envelop theorem we obtain:

$$V_M^{i,t} = -\frac{\phi'(l_t^i)}{w_t^i} = -\left(1 - \phi'(\Delta_t^i)\right) + v'(x_t^i)P'(\Delta_t^i), \quad V_B^{i,t} = 1, \quad (7)$$

combined with the equation (6), we have:

$$\phi'(\Delta_t^i) + v'(x_t^i)P'(\Delta_t^i) = T'(M_t^i), \quad (8)$$

which implies the total marginal cost of income concealment equals the marginal labor income tax at the agent's optimum.

In the Mirrlees' framework, the agent monotonicity condition is necessary. In the case of tax evasion, we follow Blomquist, Christiansen, and Micheletto (2016) and define the modified agent monotonicity condition (MAMC) as the marginal rate of substitution between reported income and net income decreases with the agent's wage rate, i.e.  $\partial MRS_{MB}^{i,t}/\partial w_t^i \equiv \partial(-V_M^{i,t}/V_B^{i,t})/\partial w_t^i$  is negative. After then, we give:

### **Proposition 1.**

- (i) *If there are no costs or only the non-monetary cost exists, then  $\Delta_w^{i,t} \geq 0$ ,  $k_w^{i,t} = 0$ .*
- (ii) *If either only the monetary cost exists or both costs exist, then  $\Delta_w^{i,t} \geq 0$ ,  $k_w^{i,t} \geq 0$ .*
- (iii) *The modified agent monotonicity condition holds in the tax evasion case, regardless of whether the costs consist of the non-monetary cost or monetary cost, or both.*

where  $(\Delta_w^{i,t}, k_w^{i,t})$  denotes the agent's behavior responses to a small change in productivity.

*Proof. See the Appendix.*

Proposition 1 implies that whether there are concealment costs or not, it is beneficial for the mimicker to hide some of his labor income. With regards to capital investment, the mimicker

would only choose to invest more than the low-skilled worker if there is a future monetary cost that he must pay.

### 3. The government's problem

The government's objective function is defined as a simple overlapping generation social welfare function where the population of each type of agent is normalized to unity:

$$\sum_{t=1}^{\infty} \beta^{t-1} \sum_{i=H,L} \theta_t^i V(B_t^i, M_t^i, \tau_t; w_t^i), \quad (9)$$

where  $\beta$  is the government's discount factor over periods and  $\theta_t^i$  ( $i = H, L$ ) represents the redistribute tastes among different types of agents.

Following Mirrlees's approach, the government designs an incentive compatibility constraint to guarantee each agent truthfully reports his income due to asymmetric information. Here we only consider the redistributive case:  $V(B_t^H, M_t^H, \tau_t; w_t^H) \geq V(B_t^L, M_t^L, \tau_t; w_t^H)$ . In addition, the government collects revenues by taxing labor incomes, and capital incomes as well as levying penalties on income concealment behavior. Then the budget constraint in period  $t$  is given by:  $\sum_{i=H,L} [M_t^i - B_t^i + \tau_t k_{t+1}^i + P(\Delta_{t-1}^i)] \geq R_t$ . Then we give the Lagrangian function:

$$\begin{aligned} \mathcal{L} = & \sum_{t=1}^{\infty} \beta^{t-1} \sum_{i=H,L} \theta_t^i V(B_t^i, M_t^i, \tau_t; w_t^i) + \sum_t \lambda_t [V(B_t^H, M_t^H, \tau_t; w_t^H) - V(B_t^L, M_t^L, \tau_t; w_t^H)] \\ & + \sum_{t=1}^{\infty} \rho_t \sum_{i=H,L} [(M_t^i - B_t^i + \tau_t k_{t+1}^i + P(\Delta_{t-1}^i)) - R_t]. \end{aligned} \quad (10)$$

Differentiate the equation (10) with respect to  $B_t^H, B_t^L, M_t^H, M_t^L$  and  $\tau_t$ , respectively, and suppose there are internal solutions, the first-order conditions are given by:

$$B_t^H: \quad (\beta^{t-1} \theta_t^H + \lambda_t) V_B^{H,t} + \rho_t \left( -1 + \tau_t \frac{\partial k_{t+1}^H}{\partial B_t^H} \right) + \rho_{t+1} P'(\Delta_t^H) \Delta_B^{H,t} = 0, \quad (11)$$

$$B_t^L: \quad \beta^{t-1} \theta_t^L V_B^{L,t} - \lambda_t \hat{V}_B^{L,t} + \rho_t \left( -1 + \tau_t \frac{\partial k_{t+1}^L}{\partial B_t^L} \right) + \rho_{t+1} P'(\Delta_t^L) \Delta_B^{L,t} = 0, \quad (12)$$

$$M_t^H: \quad (\beta^{t-1} \theta_t^H + \lambda_t) V_M^{H,t} + \rho_t \left( 1 + \tau_t \frac{\partial k_{t+1}^H}{\partial M_t^H} \right) + \rho_{t+1} P'(\Delta_t^H) \Delta_M^{H,t} = 0, \quad (13)$$

$$M_t^L: \quad \beta^{t-1} \theta_t^L V_M^{L,t} - \lambda_t \hat{V}_M^{L,t} + \rho_t \left( 1 + \tau_t \frac{\partial k_{t+1}^L}{\partial M_t^L} \right) + \rho_{t+1} P'(\Delta_t^L) \Delta_M^{L,t} = 0, \quad (14)$$

$$\begin{aligned} \tau_t: \quad & (\beta^{t-1} \theta_t^H + \lambda_t) V_{\tau}^{H,t} + \beta^{t-1} \theta_t^L V_{\tau}^{L,t} - \lambda_t \hat{V}_{\tau}^{L,t} + \rho_t \left[ \sum_{i=H,L} (\tau_t k_{t+1}^i + k_{t+1}^i) \right] \\ & + \rho_{t+1} \sum_{i=H,L} P'(\Delta_t^i) \Delta_{\tau}^{i,t+1} = 0. \end{aligned} \quad (15)$$

where the “hat” refers to the mimickers’ terms.

### 3.1. Optimal capital income taxation

Based on the proof in Appendix, the formula of optimal capital tax can be written as:

$$\tau_t = \frac{-\beta \sum_{i=H,L} P'(\Delta_t^i) \Delta_\tau^{i,t}}{\sum_{i=H,L} k_\tau^{i,t+1}} + \frac{(\gamma_t^L - \gamma_t^H)(k_{t+1}^L - \hat{k}_{t+1}^L)}{\sum_{i=H,L} k_\tau^{i,t+1}}, \quad (16)$$

where  $\gamma_t^i = \theta^i / (\theta^H + \theta^L)$  denotes the Pareto weights for different types of agents. Based on Proposition 1, it is straightforward to derive that  $(k_{t+1}^L - \hat{k}_{t+1}^L)$  takes the value of zero or a negative number, depending on the absence or presence of the monetary cost, respectively.

Then we give:

#### **Proposition 2.**

- (i) *In the absence of the monetary cost, the optimal capital taxation is zero.*
- (ii) *In cases where only the monetary cost exists or when both types of costs are present, the revenue effect indicates the need for subsidies, while the MAMC effect indicates the imposition of positive taxes.*

*Proof. See the Appendix.*

Proposition 2 implies the Atkinson-Stiglitz theorem may not hold even in the case of quasi-linear separable preference, which is no doubt due to the introduction of tax evasion concerns and the monetary cost. As the Appendix displays, tax reform in  $\tau_t$  can reduce the agent’s both hidden income and capital investment, i.e.  $\Delta_\tau^{i,t} \leq 0$ ,  $k_\tau^{i,t+1} \leq 0$ . The two terms jointly imply the government should subsidize capital investments as  $P'(\Delta_t^i)$  is positive while the denominator is negative. We name the first term on the right hand as the revenue effect. As for the second term, since the difference between  $k_{t+1}^L$  and  $\hat{k}_{t+1}^L$  is due to the modified agent monotonicity condition (MAMC) while  $(\gamma^L - \gamma^H)$  is directly dependent on how the government values the utilities between the low-skilled and the high-skilled, we then call the second term on the right hand as the MAMC effect.

### 3.2. Optimal labor income taxation

Based on the proof in the Appendix, we give the optimal labor taxation formulas:

$$T'(M_t^H) = -\tau_t k_M^{H,t+1} - \beta P'(\Delta_t^H) \Delta_M^{H,t}, \quad (17)$$

$$T'(M_t^L) = (\gamma^L - \gamma^H)(MRS_{MB}^{L,t} - \widehat{MRS}_{MB}^{L,t}) - \tau_t k_M^{L,t+1} - \beta P'(\Delta_t^L) \Delta_M^{L,t}. \quad (18)$$

We find the famous “no distortion at the top” result could no longer hold as the right-hand of the equation (17) may not be zero in the presence of the monetary cost. The first part on the right-hand side of equation (18) is standard and is a consequence of the incentive-comparability

condition. Proposition 1 states that  $(MRS_{MB}^{L,t} - \overline{MRS}_{MB}^{L,t})$  is positive, which implies positive tax rates. Besides, we can also affirm that  $-\beta P'(\Delta_t^i)\Delta_M^{i,t}$  is positive as  $\Delta_M^{i,t}$  has been proven to be negative. This finding is interesting as it suggests that the marginal tax rate should be increased to offset the loss in revenue resulting from a reduction in hidden income and, consequently, a decrease in penalty revenue. However, the  $-\tau_t k_M^{H,t+1}$  term, which represents the impact of the change in capital tax revenue, remains ambiguous due to the unknown value of  $\tau_t$ , making it impossible to determine the total effect. To sum up, we also state:

**Proposition 3.**

- (i) *In the case where only the non-monetary costs exist, tax evasion has no impact on marginal labor tax rates.*
- (ii) *However, if either the monetary cost exists or both costs are present, the following observations can be made: (a) For an agent who chooses to conceal his labor income, the assumption of tax evasion has a positive effect on his marginal tax rate through monetary cost, and has a positive or negative effect on the tax rate when  $\tau_t$  is positive or negative, respectively, through capital taxation. (b) For an agent who chooses not to conceal his income, the assumption of tax evasion does not affect their marginal tax rate.*

It comes as no surprise that the standard result regarding optimal labor taxation remains unchanged when there is no monetary cost involved in tax evasion, as demonstrated by Gahvari and Micheletto (2020). However, the introduction of the monetary cost alters the situation. Especially, if the marginal cost  $P'(\Delta_t^i)$  is sufficiently high, the impact can be significant, as reflected in the equations (16) (17), and (18).

#### 4. Conclusion

Our paper presents a novel perspective on the optimal taxation problem, highlighting two crucial elements. Firstly, we demonstrate the inadequacy of the widely accepted zero capital tax result in the presence of monetary cost and tax evasion concerns, even when considering single heterogeneity and separability. The realistic assumption that a monetary cost, or penalty, exists and arises only in the future is a crucial factor in obtaining this result. Secondly, our study offers valuable insights into the effect of various types of income concealment costs on tax schedules, highlighting the importance of monetary costs in altering the results compared to standard models without considering tax evasion. In a word, our findings offer valuable insights into addressing the optimal taxation problem in the context of tax evasion, contributing to the existing literature on the subject.

## References

- Atkinson, A. B., & Stiglitz, J. E. (1976). The design of tax structure: direct versus indirect taxation. *Journal of public Economics*, 6(1-2), 55-75.
- Blomquist, S., Christiansen, V., & Micheletto, L. (2016). Public Provision of Private Goods, Self-Selection, and Income Tax Avoidance. *The Scandinavian Journal of Economics*, 118(4), 666-692.
- Casamatta, G. (2021). Optimal income taxation with tax avoidance. *Journal of Public Economic Theory*, 23(3), 534-550.
- Christiansen, V. (1984), Which Commodity Taxes should Supplement the Income Tax? *Journal of Public Economics*, 24, 195–220.
- Christiansen, V., & Tuomala, M. (2008). On taxing capital income with income shifting. *International Tax and Public Finance*, 15(4), 527-545.
- Gahvari, F., & Micheletto, L. (2020). Wage endogeneity, tax evasion, and optimal nonlinear income taxation. *Journal of Public Economic Theory*, 22(3), 501-531.
- Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The review of economic studies*, 38(2), 175-208.
- Stern, N. H. (1982). Optimum taxation with errors in administration. *Journal of Public Economics*, 17, 181–211.
- Stiglitz, J. E. (1982). Self-selection and Pareto efficient taxation. *Journal of public economics*, 17(2), 213-240.

## Appendix

*Proof of Proposition 1.*

Partially differentiating the agent's first order conditions (2) and (3) with respect to  $B_t^i$ ,  $M_t^i$ ,  $\tau_t$ , and  $w_t^i$  ordinally we obtain  $(\Delta_B^{i,t}, k_B^{i,t})$ ,  $(\Delta_M^{i,t}, k_M^{i,t})$ ,  $(\Delta_\tau^{i,t}, k_\tau^{i,t})$  and  $(\Delta_w^{i,t}, k_w^{i,t})$ , respectively. We call them the agent's behavior responses due to the small changes in  $M_t^i$ ,  $B_t^i$ ,  $\tau_t$ , and  $w_t$  at the optimum and they are given by:

$$\Delta_B^{i,t} = k_B^{i,t} = 0, \quad (A1)$$

$$\Delta_M^{i,t} = -\frac{\varphi_{i,t}'' l_t^i}{(w_t^i)^2 (\phi_{i,t}'' + P_{i,t}'' v_{i,t}') + \varphi_{i,t}''} \leq 0, \quad (A2)$$

$$k_M^{i,t} = -\frac{P_{i,t}'}{r_t} \frac{\varphi_{i,t}'' l_t^i}{(w_t^i)^2 (\phi_{i,t}'' + P_{i,t}'' v_{i,t}') + \varphi_{i,t}''} \leq 0, \quad (A3)$$



$$\Delta_{\tau}^{i,t} = -\frac{P'_{i,t}}{r_t} \frac{(w_t^i)^2}{(w_t^i)^2(\phi''_{i,t} + P''_{i,t}v'_{i,t}) + \phi''_{i,t}} \leq 0, \quad (\text{A4})$$

$$k_{\tau}^{i,t} = \frac{1}{r_t^2 v''_{i,t}} - \left(\frac{P'_{i,t}}{r_t}\right)^2 \frac{w_t^2}{(w_t^i)^2(\phi''_{i,t} + P''_{i,t}v'_{i,t}) + \phi''_{i,t}} \leq 0, \quad (\text{A5})$$

$$\Delta_w^{i,t} = \frac{\phi'_{i,t} + \phi''_{i,t}l_t^i}{(w_t^i)^2(\phi''_{i,t} + P''_{i,t}v'_{i,t}) + \phi''_{i,t}} \geq 0, \quad (\text{A6})$$

$$k_w^{i,t} = \frac{P'_{i,t}}{r_t} \frac{\phi'_{i,t} + \phi''_{i,t}l_t^i}{(w_t^i)^2(\phi''_{i,t} + P''_{i,t}v'_{i,t}) + \phi''_{i,t}} \geq 0. \quad (\text{A7})$$

Using (6), (7) and (A6) we have

$$\frac{dMRS_{MB}^{i,t}}{dw_t^i} = -\frac{d(\varphi'(l_t^i)/w_t)}{dw_t^i} = -\frac{1}{(w_t^i)^2} [\varphi'(l_t^i) + \varphi''(l_t^i)l_t^i] + \frac{\varphi''(l_t^i)}{(w_t^i)^2} \Delta_w^{i,t} \leq 0. \quad (\text{A8})$$

*Proof of Proposition 2.*

Using Roy's identity, we have  $V_{\tau}^{i,t} = -k_{t+1}^i$  and  $\hat{V}_{\tau}^{H,t} = -\hat{k}_{t+1}^H$ . Multiplying the equation (11) and (12) by  $k_{t+1}^H$ ,  $k_{t+1}^L$  respectively and sum (15) gives:

$$\lambda_t(-k_{t+1}^L + \hat{k}_{t+1}^H) + \rho_t \tau_t \sum_{i=H,L} (k_B^{i,t+1} k_{t+1}^i + k_{\tau}^{i,t+1}) + \rho_{t+1} \sum_{i=H,L} (P'_{i,t} \Delta_B^{i,t} k_{t+1}^i + k_{\tau}^{i,t+1}) = 0. \text{ As (A1) gives } \Delta_B^{i,t} = k_B^{i,t} = 0, \text{ we further have:}$$

$$\tau_t = \frac{-\frac{\rho_{t+1}}{\rho_t} \sum_{i=H,L} N_t^i P'(\Delta_t^i) \Delta_{\tau}^{i,t} + \frac{\lambda_t}{\rho_t} (k_{t+1}^L - \hat{k}_{t+1}^L)}{\sum_{i=H,L} k_{\tau}^{i,t+1}}. \quad (\text{A9})$$

Using (11) and (12) we obtain  $\lambda_t = (\beta^{t-1}/2)(\theta^L - \theta^H)$  and  $\rho_t = (\beta^{t-1}/2)(\theta^L + \theta^H)$ , hence  $(\rho_{t+1}/\rho_t) = \beta$  and  $\lambda_t/\rho_t = (\theta^L - \theta^H)/(\theta^L + \theta^H)$ .

Only the non-monetary cost concern implies  $k_w^{i,t+1}$ ,  $P'(\cdot) = 0$ , which straightforward leads to  $\tau_t = 0$ . Similarly, only the monetary cost concern implies  $\phi(\cdot) = \phi'(\cdot) = 0$ , in which case  $k_w^{i,t+1} > 0$ , further  $k_{t+1}^L - \hat{k}_{t+1}^L < 0$  is given.

*Proof of (17) and (18).*

Using (6), (7), (11), (13), and (A1), we first have:

$$T'(M_t^H) - 1 = \frac{V_M^{H,t}}{V_B^{H,t}} = \frac{\rho_t(1 + \tau_t k_M^{H,t+1}) + \rho_{t+1} P'_{H,t} \Delta_M^{H,t}}{-\rho_t}. \quad (\text{A10})$$

Note that  $\frac{\rho_{t+1}}{\rho_t} = \beta$ , after rearrangement we then obtain (17) in the main text. In the same way, by using (6), (7), (12), (14), and (A1) we also obtain:

$$\frac{V_M^{L,t}}{V_B^{L,t}} = \frac{\lambda_t}{\rho_t} \hat{V}_B^{L,t} \left( -\frac{V_M^{L,t}}{V_B^{L,t}} + \frac{\hat{V}_M^{L,t}}{\hat{V}_B^{L,t}} \right) - (1 + \tau_t k_M^{L,t+1}) - \frac{\rho_{t+1}}{\rho_t} P'_{L,t} \Delta_M^{L,t}, \quad (\text{A11})$$

which implies (18).