ESSAYS ON ECONOMIC DYNAMICS

A Dissertation

by

SHENZHE JIANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Dennis W. Jansen
Co-Chair of Committee,	Yuzhe Zhang
Committee Members,	Sarah Zubairy
	David A. Bessler
Head of Department,	Timothy J. Gronberg

August 2017

Major Subject: Economics

ABSTRACT

The dissertation includes two sections, which apply dynamic economic models to study different economic issues.

The Section Two studies the optimal design of the Pacific Salmon Treaty, which was signed by the U.S. and Canada in 1999 to share salmon on the Pacific coast. Moral hazard exists because countries may steal from each other. If a country's observed output is suspiciously too high, the treaty either reduces the country's future share, or asks the country to make a monetary transfer to its opponent. A calibrated version of our model shows that it is optimal for the U.S. to pay Canada $327.58 million every 30.68 years. Switching to the optimal contract improves the total welfare by 1.54%.

The Section Three studies Chinese housing market.China's housing price has been growing steadily over the past decade, despite the fact that capital return has fallen dramatically. In a rational bubble framework, the fast growth rate of housing price implies a risk of the burst of housing bubble. We study the impact of bubble burst on China's economy, where the government's infrastructure investment, largely funded by land sale, is excessive. Our calibrated model shows that if the bubble bursts in 2017, then in the short run GDP growth rate falls to 2.3% due to the hit to the housing sector, but GDP in the long run exceeds that under the bubble because excessive infrastructure investment is no longer sustainable. If the bubble remains, however, implementing property tax will reduce its size and increase long-run output.

DEDICATION

To my mother, my father, and my wife.

ACKNOWLEDGMENTS

I am forever indebted to my advisor Dennis Jansen and my co-advisor Yuzhe Zhang for their invaluable guidance and help over the past few years. I am also grateful to members of my dissertation committee: Sarah Zubairy and David Bessler for their continued support and valued comments throughout the course of this research. I would also like to thank all Ph.D. students, faculty members and staff of the Economic departments at Texas A&M University.Without their help, it would be impossible for me to finish this dissertation.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Jansen [advisor], Professor Zhang [co-advisor] and Professor Zubairy of the Department of Economics and Professor Bessler of the Department of Agricultural Economics.

The analyses depicted in Section Two were conducted in part by Yuzhe Zhang of the Department of Economics.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

This dissertation was created without outside financial support.

NOMENCLATURE

IC	Incentive Compatible
ODE	Ordinary Differential Equation
PST	Pacific Salmon Treaty
OLS	Ordinary Least Squares

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

The main focus of my doctoral research is to apply different techniques in economic dynamics to study some important public policies. The first work uses dynamic mechanism design approach to study the Pacific Salmon Treaty, which was signed by the U.S. and Canada. We find switching to the optimal contract improves the total welfare by 1.54%. The second work tries to use a dynamic rational bubble model to explain the high growth of Chinese housing price and the falling capital return. We find if the asset bubble was to burst in 2017, then in the short run China's GDP would fall by 3.5 percent due to the hit to the housing sector, but GDP in the long run would exceed that under the bubble because excessive infrastructure investment would no longer be sustainable.

In the first research applies dynamic-contract method in studying Pacific Salmon Treaty. which was signed by the U.S. and Canada in 1999 to share salmon on the Pacific coast. Moral hazard exists because countries may steal from each other. If a country's observed output is suspiciously too high, the treaty either reduces the country's future share, or asks the country to make a monetary transfer to its opponent. A calibrated version of our model shows that it is optimal for the U.S. to pay Canada $327.58 million every 30.68 years. Switching to the optimal contract improves the total welfare by 1.54%. Also in this paper, we have theoretical contribution on continuous-time game theory. Sannikov (2007) shows that the boundary of the set of equilibrium payoffs satisfies a differential equation. However, the differential equation is not easy to solve because its boundary condition is unknown and needs to be obtained by trial and error. The boundary condition in our setup is given by smooth pasting conditions. Under smooth pasting conditions, we show that the optimality equation admits only two solutions. This greatly simplifies the calculation of the set of equilibrium payoffs.

In the second research, I apply asset pricing model in studying China's housing market. China's housing price has been growing fast and steadily over the past decade, despite the fact that capital return has fallen dramatically. In my job-market paper **Boom and Bust in China's Housing Market**. I build a rational bubble model to answer two important questions: 1) what is the consequence of a housing market crash? and 2) how does the adoption of property tax affect housing market? Our calibrated model shows that if the bubble was to burst in 2017, then in the short run ChinaâĂŹs GDP would fall by 3.5 percent due to the hit to the housing sector, but GDP in the long run would exceed that under the bubble because excessive infrastructure investment, which is largely funded by land sales, would no longer be sustainable. We also find that if the bubble remains, however, implementing a property tax will reduce the size of the bubble and increase long-run output.

2. ON THE PACIFIC SALMON TREATY

2.1 Introduction

Pacific salmon are a resource shared by the United States and Canada. In both countries, salmon are bred in rivers, streams, and lakes. After hatching, they go downriver to the ocean, and live there for years before returning to the freshwater habitats to spawn and die. In the ocean, salmon migrate across international boundaries, passing through coastal areas of Oregon, Washington, British Columbia, and Alaska. As a result, U.S. fishers inevitably intercept salmon originally from Canada, and vice versa.

The two countries have a long history of squabbling over their respective shares of the catch. In 1995, the state of Alaska defended its catch of sockeye (a high value species of salmon) originating in British Columbia using an argument unsupported by the United Nations Law of the Sea Convention.[1] Canada proposed third-party binding arbitration, but the U.S. opposed. In July 1997, after the Canadian government accused the U.S. of aggressive fishery, angry Canadian fishermen blockaded the Alaskan ferry *Malaspina* with 200 fishing vessels, preventing it from leaving the Prince Rupert port in British Columbia.[2] In an effort to end the escalating fish war, the two governments entered into a long-term fishing agreement under the Pacific Salmon Treaty in 1999.[3]

This paper studies the optimal design of the Pacific Salmon Treaty. It features two countries sharing a natural resource under moral hazard in an infinite-horizon model. At time zero, countries sign a contract/treaty to specify the sharing rule in all future dates.

[1] This "pasturage" argument is that salmon originating in British Columbia spend part of their life cycle within U.S. waters (see [1]). However, Article 66(1) of the United Nations Law of the Sea Convention states that "States in whose rivers anadromous species originate shall have the primary interest in and responsibility for such stocks."

[2] Another violent incident happened in May 1997, when three U.S. commercial fishing vessels were arrested for failing to comply with Canadian regulations while in Canadian waters.

[3] The original Pacific Salmon Treaty was signed in 1985, but it was ill-designed and poorly enforced. See Appendix A for a brief history of the Pacific Salmon Treaty.

Afterwards, they play a continuous-time hidden-action game *a la* [2], where the public signal is the sum of hidden action and some Brownian motion shock, and hidden action is modeled as countries deviating from the pre-specified sharing rule and stealing from each other. There are two instruments in the contract to prevent stealing. If a country's observed output is suspiciously too high or its opponent's output is suspiciously too low, the contract may (1) reduce the country's future share, and (2) ask the country to make a monetary transfer to its opponent. We allow both instruments to be fully history dependent.

Our assumptions of moral hazard and side payment are motivated by the data. First, stealing in our model corresponds to the behavior in reality that a country intercepts more salmon than what the Pacific Salmon Treaty specifies. Over interception cannot be perfectly detected because both countries' fishing weights are affected by large random shocks such as climate change.[4] Second, the U.S. has made two payments to Canada since 1999, with an average payment of $56.31 million. Since Canada's value added from salmon is only $6.16 million per year, side payments have played an important role in compensating Canada's loss so that Canada does not walk away.

There are two main results in this paper: one is theoretical and the other is empirical. Our theoretical result is on continuous-time game theory. [2] shows that the boundary of the set of equilibrium payoffs satisfies a differential equation (i.e., the optimality equation). However, the differential equation is not easy to solve because its boundary condition is unknown and needs to be obtained by trial and error. Thanks to the possibility of side payment, the boundary condition in our setup is given by smooth pasting conditions. Under smooth pasting conditions, we show that the optimality equation admits only two solutions. This greatly simplifies the calculation of the set of equilibrium payoffs.

Our empirical result recommends a policy change in the Pacific Salmon Treaty. Be-

[4]The coefficients of variation of salmon catching weights in the U.S. and Canada are, respectively, 13.37% and 43.56%.

cause the U.S. production function is more efficient than Canada's, we find that the optimal contract assigns a bigger share of salmon to the U.S. than the current treaty does. Making this change will improve the total welfare of both countries by 1.54%.

Although we only study the sharing of Pacific salmon between the U.S. and Canada in the paper, international fishery sharing agreements are actually very common. In the database of the Food and Agriculture Organization of the United Nations, there are 1927 agreements registered (over 300 of which are signed by the U.S.) and the earliest can be traced back to the year 1351. Many agreements involve disputes. For instance, disputes have occurred multiple times between Australia and New Zealand in the South Tasman Rise Trawl fishery, and also between the United Kingdom and Iceland in the North Atlantic cod fishery. Moreover, the issue of sharing natural resources with side payment goes beyond fisheries. In an international river-sharing agreement, the upstream country of the Syr Darya River, Kyrgyzstan, agreed to increase summer discharges to supply to the downstream country, Uzbekistan, in exchange for fossil fuel transfers.

Related literature There are two theories of incentives in the literature on dynamic games and contracts. One emphasizes variations of continuation payoff as an incentive device and the other emphasizes side payments. [3] and [2] develop methods for solving moral-hazard repeated games, in discrete time and continuous time, respectively. In these games, the only incentive is provided by the sensitivity of continuation payoffs to signals about players' actions. Moreover, the folk theorem states that such incentive is strong enough to support any individual-rational payoff as an equilibrium outcome, when players are patient. On the other hand, recent papers on relational contracts (e.g., [4], [5], [6], [7]) emphasize the role played by side payments. They show that, when people are risk neutral and side payments are possible, optimal contracts (or efficient equilibria) are stationary, i.e., the same payment scheme and action profile are repeated in every period. In stationary contracts (or equilibria), side payments completely crowd out variation of continuation payoffs as an incentive device, because continuation payoffs are fixed. Our paper bridges the above two theories by incorporating side payments into a continuous-time game built on [2]. Both incentives are utilized in our model: continuation payoffs vary with public signals at all times, while side payments are used only when one player's participation constraint binds. In contrast to relational contracts where people "settle up" with side payments each period, payments are used less frequently in our model because payments are costless in relational contracts, but costly here.

Our paper is related to a large literature on the tragedy of the commons (e.g., [8], [9], [10], [11], [12], [13]). When a resource is shared by many individuals, a tragedy of the commons occurs because an individual does not internalize the effect that his or her extraction reduces future stock of the resource, and hence reduces the welfare of other individuals. Consequently, the equilibrium level of extraction is more than what is socially optimal. Our paper differs from this literature because we assume a fixed stock of salmon over time, and thus the issue of over-extraction does not exist. This simplifying assumption is supported by the fact that the salmon stock has been well preserved since 1999 (see the last paragraph of Appendix A). Another difference is that while the tragedy-of-the-commons literature typically assumes away hidden action/moral hazard, hidden action plays an important role in our model: without it, side payments are no longer needed as an incentive device in our optimal contract.

Our paper is organized as follows. Section 2.2 studies the model under linear cost for side payment. The main result is that the optimality equation admits two solutions. Section 2.3 extends the model to the case with fixed cost for side payment. We calibrate the fixed-cost model and find that the welfare gain of switching to the optimal contract is 1.54%. Section 2.4 concludes. Appendix A provides a brief history of the Pacific Salmon Treaty, Appendix B has additional calibration details, and Appendix C contains proofs of all the results.

2.2 Model

Two players share a natural resource through long-term contracting at time zero. If the players sign a contract, then after time zero they participate in a repeated game with moral hazard in continuous time; otherwise, they receive outside options $\underline{u}_1$ and $\underline{u}_2$, respectively. Before solving the contracting problem at time zero, we shall first describe the details of the repeated game after time zero.

At each time $t \in [0, \infty)$, there is one unit of natural resource to be shared. If player 1 gets share $x \in [0, 1]$, then player 2 gets $1 - x$. The players' payoffs are, respectively, $p_1(x)$ and $p_2(x)$, where p_1 is strictly increasing and concave in x, and p_2 is strictly decreasing and concave in x.

Moral hazard exists because players may steal the natural resource from their opponents. If player i steals, then player j loses $1 + \mu$ dollars whenever player i gains 1 dollar. We can interpret $\mu > 0$ as additional damages to player j's environment caused by illegal extraction. If the two players' payoffs are $(p_1(x_t), p_2(x_t))$ and their stealing efforts are (e_{1t}, e_{2t}), then their utility flows are:

$$(p_1(x_t) + e_{1t})dt + (-(1 + \mu)e_{2t}dt + \sigma_1 dZ_{1t}),$$
$$(p_2(x_t) + e_{2t})dt + (-(1 + \mu)e_{1t}dt + \sigma_2 dZ_{2t}).$$

Here, Z_{1t} and Z_{2t} are two independent standard Brownian motions that represent the uncertainty in the environment, and (σ_1, σ_2) measure the size of uncertainty.[5]

Players' actions (e_{1t}, e_{2t}) are private (i.e., e_{it} is not observable by player j), but their utility flows are public. The public history at time t contains the observed utility flows

[5] We have modeled stealing as players stealing each other's payoffs rather than the underlying resource. A more realistic model would specify players' utilities as $p_1(x_t + e_{1t} - (1+\mu)e_{2t})dt + \sigma_1 dZ_{1t}$ and $p_2(x_t + (1 + \mu)e_{1t} - e_{2t})dt + \sigma_2 dZ_{2t}$. Our specification makes it easier to calibrate μ in Section 2.3.2, because financial gains and losses due to illegal fishing are reported in dollar amounts, not in physical units of the resource (such as tons in the fishing industry).

before t and is captured by the filtration $\{\mathcal{F}_t\}$. Because player 1 knows both his utility flow and his action e_{1t}, he can infer $(-(1+\mu)e_{2t}dt + \sigma_1 dZ_{1t})$, but he cannot distinguish $-(1+\mu)e_{2t}dt$ from the uncertainty $\sigma_1 dZ_{1t}$. In other words, if player 1 experiences a year of low output, he does not know whether it is due to player 2's stealing, or unfavorable climate change in that year.

We allow for side payments between the two players. In particular, if player i pays q_i to player j, then the total expense for player i is $(1+\tau)q_i$, where $\tau > 0$ represents a linear transaction cost. Player i's discounted payoff is

$$\begin{aligned} W_i &= E\left[\int_0^\infty re^{-rt}\left((p_{it} + e_{it} - (1+\mu)e_{jt} - (1+\tau)q_{it} + q_{jt})dt + \sigma_i dZ_{it}\right)\right] \quad (2.1) \\ &= E\left[\int_0^\infty re^{-rt}\left((p_{it} + e_{it} - (1+\mu)e_{jt})dt - (1+\tau)dQ_{it} + dQ_{jt} + \sigma_i dZ_{it}\right)\right], \end{aligned}$$

where $r > 0$ is the discount rate and $Q_{it} := \int_0^t q_{is}ds$ is the cumulative payment by player i up to time t.

A long-term contract that the two players enter at time zero specifies a sharing-rule process $\mathbf{x} = \{x_t; t \geq 0\}$, action processes $\mathbf{e} = \{(e_{1t}, e_{2t}); t \geq 0\}$ for the players to take, and side-payment processes $\mathbf{Q} = \{(Q_{1t}, Q_{2t}); t \geq 0\}$. Processes $(\mathbf{x}, \mathbf{e}, \mathbf{Q})$ must be adapted to the public information available to the players, $\{\mathcal{F}_t\}$.

2.2.1 Incentive compatibility

Because actions (e_{1t}, e_{2t}) are not publicly observable, contracts will have to satisfy incentive compatibility constraints. A contract is incentive compatible (IC) if stealing cannot make the players better off. We will express IC constraints using the following results of [2].

For any contract $(\mathbf{x}, \mathbf{e}, \mathbf{Q})$, define its associated continuation payoff process $\mathbf{W} =$

$\{(W_{1t}, W_{2t}); t \geq 0\}$ as

$$W_{it} := E\Big[\int_t^\infty re^{-r(s-t)}\Big((p_{is} + e_{is} - (1+\mu)e_{js})ds - (1+\tau)dQ_{is} + dQ_{js} + \sigma_i dZ_{is}\Big)\Big]. \tag{2.2}$$

There exists a process $\beta^i = \{(\beta_t^{i1}, \beta_t^{i2}); t \geq 0\}$ such that the continuation payoffs evolve as

$$\begin{aligned} dW_{it} &= r(W_{it} - (p_{it} + e_{it} - (1+\mu)e_{jt}))dt + r(1+\tau)dQ_{it} - rdQ_{jt} \\ &\quad + r\beta_t^{i1}dZ_{1t} + r\beta_t^{i2}dZ_{2t}, \quad i = 1, 2. \end{aligned}$$

Here, β_t^{ij} represents the sensitivity of player i's continuation payoff W_{it} to player j's signal (i.e., player j's observed output). The IC constraint for player i is, for all $t \geq 0$ and $\tilde{e}_i \geq 0$,

$$e_{it} + \beta_t^{ii}e_{it} - (1+\mu)\beta_t^{ij}e_{it} \geq \tilde{e}_i + \beta_t^{ii}\tilde{e}_i - (1+\mu)\beta_t^{ij}\tilde{e}_i.$$

This means

$$1 + \beta_t^{ii} - (1+\mu)\beta_t^{ij} \leq 0, \tag{2.3}$$

with equality if $e_{it} > 0$.

2.2.2 The set of continuation payoffs

Since players have outside options $(\underline{u}_1, \underline{u}_2)$, their continuation payoffs in any contract must satisfy the participation constraint:

$$W_{it} \geq \underline{u}_i, \quad \forall t \geq 0, i = 1, 2. \tag{2.4}$$

For convenience, we can define $\mathcal{A} := \{(w_1, w_2) : w_1 \geq \underline{u}_1, w_2 \geq \underline{u}_2\}$ and write the participation constraint as $(W_{1t}, W_{2t}) \in \mathcal{A}$.

Let $\mathcal{E}$ be the set of continuation payoff pairs achieved by all IC contracts, i.e.,

$$\mathcal{E} := \{(W_{1t}, W_{2t}) \in \mathcal{A} : (W_{1t}, W_{2t}) \text{ satisfies (2.2) for some IC contract}\}.$$

The following lemma shows that $\mathcal{E}$ is monotonic: if (w_1, w_2) belongs to $\mathcal{E}$, then the vectors below (w_1, w_2) also belong to $\mathcal{E}$. In particular, $(\underline{u}_1, \underline{u}_2) \in \mathcal{E}$.

Lemma 1. *If $(w_1, w_2) \in \mathcal{E}$ and for some $Q_1 \geq 0$ and $Q_2 \geq 0$,*

$$\begin{aligned} \tilde{w}_1 &= w_1 - r(1+\tau)Q_1 + rQ_2 \geq \underline{u}_1, \\ \tilde{w}_2 &= w_2 - r(1+\tau)Q_2 + rQ_1 \geq \underline{u}_2, \end{aligned}$$

then $(\tilde{w}_1, \tilde{w}_2) \in \mathcal{E}$.

The proof of this lemma is straightforward. Suppose at time 0 a contract lets player i pay Q_i to his opponent and restarts from (w_1, w_2). This contract delivers payoff $w_1 - r(1+\tau)Q_1 + rQ_2$ to player 1 and $w_2 - r(1+\tau)Q_2 + rQ_1$ to player 2, thus $(\tilde{w}_1, \tilde{w}_2) \in \mathcal{E}$.

The next subsection studies the boundary of $\mathcal{E}$, which we denote as $\partial\mathcal{E}$.

2.2.3 Pareto frontier

This subsection shows that $\partial\mathcal{E}$ consists of three portions: a horizontal portion, a vertical portion, and a downward sloping portion.

We first characterize the horizontal portion and the vertical portion of $\partial\mathcal{E}$. Define

$$\overline{u}_1 := \max\{w_1 : (w_1, \underline{u}_2) \in \mathcal{E}\}, \quad \overline{u}_2 := \max\{w_2 : (\underline{u}_1, w_2) \in \mathcal{E}\}. \tag{2.5}$$

That is, $\overline{u}_i$ is the highest payoff for player i when player j's payoff is at the minimum. Then the horizontal portion and vertical portion of $\partial\mathcal{E}$ are, respectively, $\{(w_1, \underline{u}_2) : \underline{u}_1 \leq w_1 \leq \overline{u}_1\}$ and $\{(\underline{u}_1, w_2) : \underline{u}_2 \leq w_2 \leq \overline{u}_2\}$. To see that these boundaries are non-degenerate, we need to show $\underline{u}_i < \overline{u}_i$. Pick $w \in \mathcal{E}$ such that $w \neq (\underline{u}_1, \underline{u}_2)$. Lemma 1 states that if $Q_1 = \frac{w_1 - \underline{u}_1}{r(1+\tau)}$ and $Q_2 = 0$, then $(\tilde{w}_1, \tilde{w}_2) = (\underline{u}_1, w_2 + (w_1 - \underline{u}_1)/(1+\tau)) \in \mathcal{E}$. Therefore,

$$\underline{u}_2 < w_2 + (w_1 - \underline{u}_1)/(1+\tau) \leq \overline{u}_2.$$

Similarly, we can show that $\underline{u}_1 < \overline{u}_1$.

Then, we study the portion of the boundary from $(\overline{u}_1, \underline{u}_2)$ to $(\underline{u}_1, \overline{u}_2)$. This portion stays above the straight line connecting the two points because $\mathcal{E}$ is convex. We show below that this boundary is downward sloping.

Let $\mathbf{T}(w)$ and $\mathbf{N}(w)$ be the unit tangent and outward normal vectors to $\partial\mathcal{E}$ at a boundary point w. If θ is the angle between $\mathbf{N}(w)$ and the x-axis, then $\mathbf{T}(w) = (-\sin(\theta), \cos(\theta))$ and $\mathbf{N}(w) = (\cos(\theta), \sin(\theta))$. To see that the boundary from $(\overline{u}_1, \underline{u}_2)$ to $(\underline{u}_1, \overline{u}_2)$ is downward sloping, it is sufficient to show that the slope of the boundary, $-\frac{\cos(\theta)}{\sin(\theta)}$, satisfies

$$-(1+\tau) \leq -\frac{\cos(\theta)}{\sin(\theta)} \leq -(1+\tau)^{-1}. \tag{2.6}$$

Lemma 1 states that $(w_1 - r(1+\tau)Q_1, w_2 + rQ_1) \in \mathcal{E}$, $(w_1 + rQ_2, w_2 - r(1+\tau)Q_2) \in \mathcal{E}$ for $Q_1 \in [0, \frac{w_1 - \underline{u}_1}{r(1+\tau)}]$ and $Q_2 \in [0, \frac{w_2 - \underline{u}_2}{r(1+\tau)}]$. Since the convex set $\mathcal{E}$ is either above or below the tangent line at w, we know that $(w_1 - r(1+\tau)Q_1, w_2 + rQ_1)$ and $(w_1 + rQ_2, w_2 - r(1+\tau)Q_2)$ stay on the same side of the tangent line. This implies (2.6).

Since the boundary from $(\overline{u}_1, \underline{u}_2)$ to $(\underline{u}_1, \overline{u}_2)$ is downward sloping, we can define the mapping from w_1 to w_2 as $f(\cdot)$. Function $f(\cdot)$ represents the Pareto frontier of $\mathcal{E}$ because $w_2 = f(w_1)$ is the highest payoff for player 2 if player 1's payoff is $w_1 \in [\underline{u}_1, \overline{u}_1]$.

In the following analysis, we further characterize the Pareto frontier as the solution to some differential equation.

2.2.4 Restriction to no stealing

In this paper, we restrict ourselves to contracts that recommend no stealing, i.e., $e_{1t} = e_{2t} = 0$ for all t. Under the following assumption, this restriction is without loss of generality.

Assumption 1. *Stealing is more costly than making a transfer through side payments, i.e.,* $\mu \geq \tau$.

Lemma 2. *Under Assumption 1, it is optimal to recommend* $e_{1t} = e_{2t} = 0$ *for all* t.

Two remarks regarding the assumption are in order. First, data suggest that stealing is costly. In Section 2.3.2, we calibrate μ to be 0.84, which seems to be much larger than typical costs associated with money transfer. Second, even if Assumption 1 is violated, a treaty that allows countries to steal might be too controversial to be politically viable.

2.2.5 Optimality equation for the Pareto frontier

[2] shows that, at any point w on the Pareto frontier, the sensitivities $\beta^i = (\beta^{i1}, \beta^{i2})$ in the IC constraints are given by a vector of tangential volatilities (ϕ_1, ϕ_2) as follows:

$$\begin{bmatrix} \beta^{11} & \beta^{12} \\ \beta^{21} & \beta^{22} \end{bmatrix} = \begin{bmatrix} -\sin(\theta)\phi_1 & \sin(\theta)\phi_2 \\ \cos(\theta)\phi_1 & -\cos(\theta)\phi_2 \end{bmatrix}. \tag{2.7}$$

Furthermore, the curvature of the boundary at w satisfies the following *optimality equation*:

$$\kappa(w) = \max_{\substack{x\in[0,1], \\ q_1\geq 0, q_2\geq 0}} \frac{2\mathbf{N}(\theta)((p_1(x), p_2(x)) + q_1(-1-\tau, 1) + q_2(1, -1-\tau) - w)}{r|\phi(\theta)|^2},$$

where $|\phi(\theta)|^2$ is given by

$$\begin{aligned} |\phi(\theta)|^2 = \min_{\phi_1,\phi_2} \quad & \sigma_1^2\phi_1^2 + \sigma_2^2\phi_2^2 \\ \text{subject to} \quad & (2.3), (2.7). \end{aligned}$$

Lemma 3 below shows that both $|\phi(\theta)|^2$ and the optimal x in the optimality equation can be solved explicitly. This result greatly simplifies the analysis in the next subsection.

Lemma 3. *1. The optimal (ϕ_1, ϕ_2) is given by*

$$(\phi_1,\phi_2) = \begin{cases} \left(\frac{\sigma_2^2}{\sin(\theta)(\sigma_2^2+(1+\mu)^2\sigma_1^2)}, \frac{(1+\mu)\sigma_1^2}{\sin(\theta)(\sigma_2^2+(1+\mu)^2\sigma_1^2)}\right), & \text{if } \theta \in (0,\theta_1]; \\ \left(\frac{\frac{1+\mu}{\cos(\theta)}-\frac{1}{\sin(\theta)}}{(1+\mu)^2-1}, \frac{\frac{1+\mu}{\sin(\theta)}-\frac{1}{\cos(\theta)}}{(1+\mu)^2-1}\right), & \text{if } \theta \in [\theta_1,\theta_2]; \\ \left(\frac{(1+\mu)\sigma_2^2}{\cos(\theta)(\sigma_1^2+(1+\mu)^2\sigma_2^2)}, \frac{\sigma_1^2}{\cos(\theta)(\sigma_1^2+(1+\mu)^2\sigma_2^2)}\right), & \text{if } \theta \in [\theta_2, \frac{\pi}{2}), \end{cases}$$

where $\theta_1 := \arctan\left(\frac{(\sigma_1^2+\sigma_2^2)(1+\mu)}{\sigma_2^2+(1+\mu)^2\sigma_1^2}\right)$ *and* $\theta_2 := \arctan\left(\frac{\sigma_1^2+(1+\mu)^2\sigma_2^2}{(\sigma_1^2+\sigma_2^2)(1+\mu)}\right)$.

2. The optimal x^* is uniquely pinned down by $\cos(\theta)p_1'(x^*)+\sin(\theta)p_2'(x^*) = 0$ for any θ.

Remark 1. *Both ϕ_1 and ϕ_2 are positive in Lemma 3. This and (2.7) imply that $\beta^{ii} < 0$ but $\beta^{ji} > 0$. That is, when player i's output is higher than the expectation ($dZ_{it} > 0$), then player i is punished (both his continuation value W_{it} and his future share of salmon decrease), while player j is rewarded. This property of the model is consistent with what we observe in the data. In 2001, the Pacific Salmon Treaty reallocated 57,000 sockeye salmon in Fraser River (about 3.7% of the total allowable catch) from the U.S. to Canada, because the U.S. catch was excessive in 2000.*[6]

We can simplify the right-hand side of the optimality equation by removing q_1 and q_2

[6]See the 17th Annual Report of the Pacific Salmon Commission.

from the numerator. Since θ satisfies (2.6),

$$\begin{aligned}\mathbf{N}(\theta)(-1-\tau,1) &= -(1+\tau)\cos(\theta)+\sin(\theta) \le 0,\\ \mathbf{N}(\theta)(1,-1-\tau) &= \cos(\theta)-(1+\tau)\sin(\theta) \le 0.\end{aligned}$$

Therefore, the optimal solutions in the maximization problems $\max_{q_1\ge 0}\mathbf{N}(\theta)(-1-\tau,1)q_1$ and $\max_{q_2\ge 0}\mathbf{N}(\theta)(1,-1-\tau)q_2$ can be set as $q_1 = 0$ and $q_2 = 0$.[7] Then the optimality equation reduces to

$$\kappa(w) = \max_{x\in[0,1]} \frac{2\mathbf{N}(\theta)((p_1(x),p_2(x))-w)}{r|\phi(\theta)|^2}.$$

2.2.6 Solving the optimality equation

This subsection characterizes solutions to the optimality equation.

On the boundary, the Pareto frontier must satisfy the smooth pasting conditions:

$$\begin{aligned}f'(w_1=\underline{u}_1) &= -(1+\tau)^{-1},\\ f'(w_1=\overline{u}_1) &= -(1+\tau).\end{aligned}$$

Smooth pasting conditions are imposed because w_{1t} is a regulated diffusion process that stays within the domain $[\underline{u}_1,\overline{u}_1]$: if $w_{1t}\in(\underline{u}_1,\overline{u}_1)$, we have argued earlier that side payments (q_1,q_2) are not used; if w_{1t} reaches either $\underline{u}_1$ or $\overline{u}_1$, side payments are just enough to keep w_{1t} inside the interval $[\underline{u}_1,\overline{u}_1]$.[8] [9] Because $f'(w_1) = -\frac{1}{\tan(\theta)}$, the smooth pasting

[7]This argument applies only in the interior of the Pareto frontier, i.e., when $w_1 > \underline{u}_1$ and $w_2 > \underline{u}_2$. If $w_i = \underline{u}_i$, then player i must pay to avoid his participation constraint being violated.

[8]See chapter 10 in [14] for a detailed discussion of regulated process. We have proven rigorously that side payments in the optimal contract are just enough to keep w_{1t} inside the interval $[\underline{u}_1,\overline{u}_1]$. This proof is available upon request.

[9]Because $w_{2t} = f(w_{1t})$ and w_{1t} is regulated to stay in $[\underline{u}_1,\overline{u}_1]$ at all times, our continuation payoff pair (w_{1t},w_{2t}) is always on the Pareto frontier, meaning that our contract is renegotiation-proof. In contrast,

conditions can be expressed in terms of θ:

$$\theta = \overline{\theta} := \arctan(1+\tau), \qquad \text{if } w_1 = \underline{u}_1, \tag{2.8}$$

$$\theta = \underline{\theta} := \arctan((1+\tau)^{-1}), \quad \text{if } w_1 = \overline{u}_1. \tag{2.9}$$

If l denotes the distance of the Pareto frontier from $(\overline{u}_1, \underline{u}_2)$ to w, then we can reformulate the optimality equation by writing all other variables as functions of l:

$$\frac{dw_1(l)}{dl} = -\sin(\theta(l)), \tag{2.10}$$

$$\frac{dw_2(l)}{dl} = \cos(\theta(l)), \tag{2.11}$$

$$\frac{d\theta(l)}{dl} = \max_{x\in[0,1]} \frac{2\mathbf{N}(\theta)((p_1(x), p_2(x)) - w)}{r|\phi(\theta)|^2}. \tag{2.12}$$

In the rest of this section, we shall characterize the solutions to (2.10)-(2.12) under the smooth pasting conditions.[10] Our main result is that the system admits two solutions, and the outer solution is the true Pareto frontier. This result greatly simplifies the search for the Pareto frontier.

In particular, the construction of a solution takes the following three steps. First, we guess a value of $\overline{u}_1$ and then solve (2.10)-(2.12) using the initial conditions $(w_1, w_2, \theta) = (\overline{u}_1, \underline{u}_2, \underline{\theta})$. Second, we show that there exists L such that $w_1(L) = \underline{u}_1$. In other words, the solution curve will cross the vertical straight line $\mathbf{Y} := \{(\underline{u}_1, w_2) : w_2 \in \mathbb{R}\}$ at some point. Third, we check the smooth pasting condition (2.8) at $l = L$, i.e., $\theta(L) = \overline{\theta}$. We can show that there are only two values of $\overline{u}_1$ starting from which $\theta(L) = \overline{\theta}$.

equilibria in [2] are not necessarily renegotiation-proof because his payoff pair may move to a boundary point below the Pareto frontier.

[10]Note that the cost of side payments, τ, affects the boundary conditions of (2.10)-(2.12), although τ does not enter the system explicitly. We can show that a lower τ moves the Pareto frontier up and to the right, and hence the set of continuation payoffs expands with better side-payment technology. This is consistent with findings in [15], [16], and [17] that side payments could make collusion/cooperation easier.

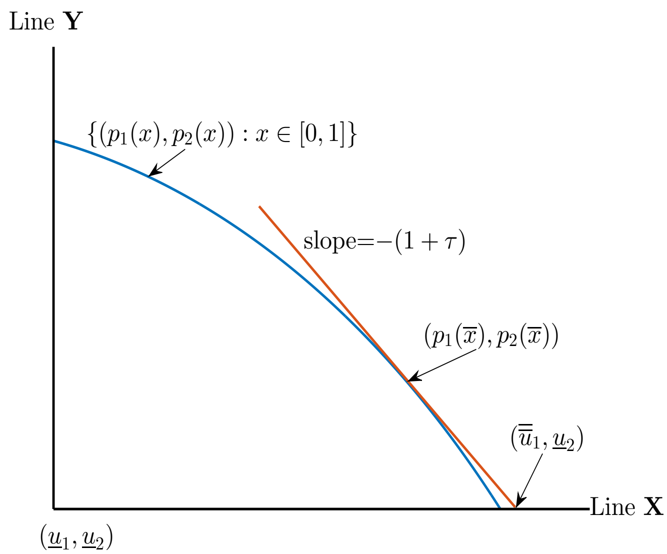

Figure 2.1: $\overline{\overline{u}}_1$ is an upper bound for $\overline{u}_1$.

In order to search for $\overline{u}_1$, we first need to find its range. The following lemma provides an upper bound for $\overline{u}_1$. Consider the tangent line of slope $-(1+\tau)$ in Figure 2.1. This line is tangential to the curve $\{(p_1(x), p_2(x)) : x \in [0,1]\}$ at point $(p_1(\overline{x}), p_2(\overline{x}))$ and intersects the horizontal straight line $\mathbf{X} := \{(w_1, \underline{u}_2) : w_1 \in \mathbb{R}\}$ at $(\overline{\overline{u}}_1 := p_1(\overline{x}) + (p_2(\overline{x}) - \underline{u}_2)/(1+\tau), \underline{u}_2)$.[11] The following lemma shows that our guess of $\overline{u}_1$ cannot exceed $\overline{\overline{u}}_1$.

Lemma 4. *If* (W_1, W_2) *is the pair of promised payoffs delivered by some contract and* $W_2 = \underline{u}_2$*, then* $W_1 \leq \overline{\overline{u}}_1$.

By Lemma 4, the appropriate range of $\overline{u}_1$ is $[\underline{u}_1, \overline{\overline{u}}_1]$. When $\overline{u}_1$ is in this range, the solution to (2.10)-(2.12) starting from $(w_1, w_2, \theta) = (\overline{u}_1, \underline{u}_2, \underline{\theta})$ exists and is unique, because the right-hand sides of (2.10)-(2.12) are Lipschitz continuous (Lemma C.1 in Appendix C

[11] $\overline{x}$ is given by the condition

$$\frac{p_2'(\overline{x})}{p_1'(\overline{x})} = -(1+\tau).$$

verifies the Lipschitz continuity).

The next lemma shows that the solution to (2.10)-(2.12) always crosses the vertical straight line **Y**. This result is not surprising because the Pareto frontier is concave, that is, the frontier becomes flatter as w_1 decreases (see Lemma C.2 in Appendix C for a proof). Intuitively, if the solution to (2.10)-(2.12) never hits the vertical **Y**, it must bend upward as w_1 approaches $\underline{u}_1$.

Lemma 5. *The solution curve starting from* $(w_1, w_2, \theta) = (\overline{u}_1, \underline{u}_2, \underline{\theta})$ *crosses **Y** once for* $\overline{u}_1 \in [\underline{u}_1, \overline{\overline{u}}_1]$. *In other words, there exists a unique* $L \geq 0$ *such that* $w_1(L) = \underline{u}_1$.

If a curve starts from $(w_1, w_2, \theta) = (\overline{u}_1, \underline{u}_2, \underline{\theta})$, denote the angle of the curve when it crosses **Y**, $\theta(L)$, as $\Theta(\overline{u}_1)$. Theorem 1 shows that $\Theta(\overline{u}_1)$ is first increasing and then decreasing in $\overline{u}_1$.

Theorem 1. $\Theta(\overline{u}_1)$ *is single-peaked. That is, there exists a* $\overline{u}_1^*$ *such that* Θ *is strictly increasing in* $[\underline{u}_1, \overline{u}_1^*]$ *and strictly decreasing in* $[\overline{u}_1^*, \overline{\overline{u}}_1]$.

The intuition for the single-peakedness in Theorem 1 is as follows. We can express $\Theta(\overline{u}_1)$ as

$$\Theta(\overline{u}_1) = \theta(L) = \underline{\theta} + \int_0^L \theta'(l)dl = \underline{\theta} + L\kappa,$$

where κ is the average curvature of the solution curve. The length of the Pareto frontier, L, is increasing in $\overline{u}_1$ while the curvature κ is decreasing in $\overline{u}_1$: as $\overline{u}_1$ increases, the Pareto frontier moves outward and its length L increases, but the larger payoff pair w reduces the term $((p_1(x), p_2(x)) - w)$ in the optimality equation and hence reduces the curvature of the Pareto frontier (see Figure 2.2). Moreover, if $\overline{u}_1$ takes the smallest value of $\underline{u}_1$, then $L = 0$ because the starting point of the curve is on **Y** and the curve is degenerate; if $\overline{u}_1$

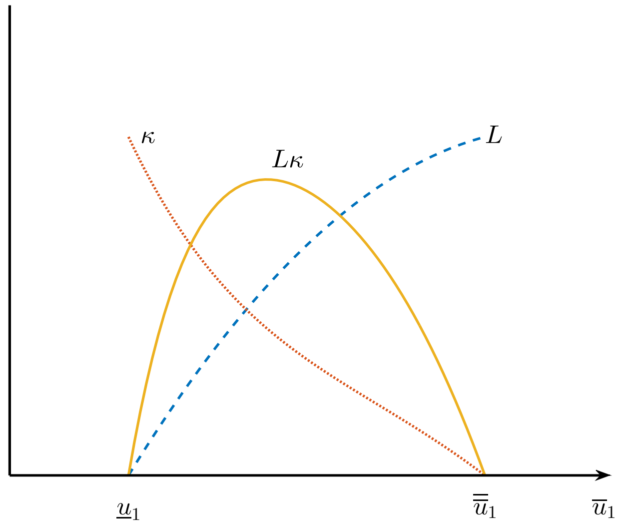

Figure 2.2: $L\kappa$ is first increasing and then decreasing in $\overline{u}_1$.

takes the largest value of $\overline{\overline{u}}_1$, then the solution curve is a straight line: its curvature is zero because $\max_{x\in[0,1]} \mathbf{N}(\theta)((p_1(x), p_2(x)) - w)$ is everywhere zero.

Because L and κ move in opposite directions and $\Theta(\overline{u}_1)$ depends on the product of the two, the slope of Θ depends on whether L or κ moves faster, in percentage terms. If $\overline{u}_1 \approx \underline{u}_1$, then L moves faster in percentage terms than κ because L is close to zero. On the other hand, κ moves faster when $\overline{u}_1 \approx \overline{\overline{u}}_1$ and κ is close to zero. Therefore, Θ is first increasing when L dominates and then decreasing when κ dominates.

Theorem 1 immediately implies the following:

Corollary 1. *Only two solutions to* (2.10)-(2.12) *satisfy the smooth pasting conditions* (2.8)-(2.9).

It is useful to relate our result to the repeated-game literature. In discrete-time repeated games, [3] show that the equilibrium set of continuation payoffs, $\mathcal{E}$, is a fixed point of some operator that maps the set of future continuation payoffs into the set of current continuation payoffs. To compute $\mathcal{E}$, they start with a superset of $\mathcal{E}$ and iterate on a sequence of sets until the sequence converges to $\mathcal{E}$. This iteration procedure, however, is time consuming. In a continuous-time setting, [2] shows that the boundary of $\mathcal{E}$ solves a second-order differential equation (i.e., the optimality equation). Sannikov's continuous-time method has an advantage over the discrete-time method because solving differential equations is numerically easier than computing the operator in [3]. The disadvantage, however, is that finding the initial condition for the optimality equation is difficult. When there are multiple initial conditions (starting from which the solution curve forms the boundary of some set), it is hard to tell which initial condition is correct.[12] Presumably, we need to find the largest set whose boundary solves the optimality equation, but this process is one of trial and error.

Our paper moves forward the analysis in [2] in two aspects. First, our initial condition is given by the smooth pasting conditions, thanks to the technology of side payments in our model. Second, under the smooth pasting conditions, only two solutions exist for the optimality equation. The two solutions can be easily distinguished because $\Theta'(\overline{u}_1)$ is positive for the inner solution but negative for the outer solution. The latter represents the true Pareto frontier because it is the boundary of a bigger set.

2.3 Model with fixed cost

Despite being analytically tractable, the linear-cost model in Section 2.2 has one drawback: it predicts that payments are made infinitely many times and each payment is infinitesimal.[13] This prediction is inconsistent with the data, because the actual payments in the Pacific Salmon Treaty are both infrequent and large (more details are in Section 2.3.2).

[12]To make matters worse, the number of initial conditions is typically unknown *a priori*.

[13]Payments are made whenever W_{1t} reaches the boundary of $[\underline{u}_1, \overline{u}_1]$. Since W_{1t} is a diffusion process, on the one hand it reaches the boundary of $[\underline{u}_1, \overline{u}_1]$ infinitely many times, but on the other hand it stays at the boundary for zero amount of time. It is merely an artifact of the continuous-time setup that infinitely many payments can be made in no time.

To better capture the payment frequency and amount, this section studies a model where payment incurs a fixed cost $C > 0$. In particular, if player i makes a payment Q to player j, then their payoffs are $-Q - \frac{C}{2}$ and $Q - \frac{C}{2}$, respectively. Our assumption that players equally share the fixed cost is convenient but not essential for the following analysis. Facing the fixed cost, players pay only occasionally in the optimal contract.

2.3.1 Pareto frontier

Our focus is still on the equilibrium set of payoffs $\mathcal{E}$ and its boundary $\partial\mathcal{E}$. Similar to the linear-cost case, $\partial\mathcal{E}$ contains a horizontal portion, a vertical portion, and a downward sloping portion (see Lemma C.7 in Appendix C for a proof).[14] We continue to denote the Pareto frontier as $f : [\underline{u}_1, \overline{u}_1] \to w_2$. Similar to the linear-cost case, side payments are not used unless w_1 is on the boundary of $[\underline{u}_1, \overline{u}_1]$ (see Lemma C.6 in Appendix C for a proof). Therefore, the interior of the Pareto frontier still satisfies the ODE in (2.10)-(2.12), which is the optimality equation with no side payments.

Boundary conditions for the ODE in (2.10)-(2.12) are given by the following value matching conditions:

$$\overline{u}_1 + \underline{u}_2 = \max_{(w_1,w_2)\in\mathcal{E}} w_1 + w_2 - rC, \tag{2.13}$$

$$\underline{u}_1 + \overline{u}_2 = \max_{(w_1,w_2)\in\mathcal{E}} w_1 + w_2 - rC. \tag{2.14}$$

We prove (2.13) in two steps. First, we show $\overline{u}_1 + \underline{u}_2 \geq \max_{(w_1,w_2)\in\mathcal{E}} w_1 + w_2 - rC$. Let (w_1^*, w_2^*) be the solution to $\max_{(w_1,w_2)\in\mathcal{E}} w_1 + w_2$. Then $(w_1^* + w_2^* - rC - \underline{u}_2, \underline{u}_2) \in \mathcal{E}$, because a contract that first lets player 2 pay $Q_2 = (w_2^* - \underline{u}_2)/r - C/2$ to player 1 and then restarts at (w_1^*, w_2^*) delivers payoff $w_2^* + r(-Q_2 - C/2) = \underline{u}_2$ to player 2 and $w_1^* + r(Q_2 - C/2) = w_1^* + w_2^* - rC - \underline{u}_2$ to player 1. Since $\overline{u}_1$ is the highest payoff for player 1 when player 2's payoff is $\underline{u}_2$, $\overline{u}_1 \geq w_1^* + w_2^* - rC - \underline{u}_2$. Second, we show

[14] The only exception is that C is so large that $\mathcal{E}$ becomes degenerate.

$\overline{u}_1 + \underline{u}_2 \leq \max_{(w_1,w_2)\in\mathcal{E}} w_1 + w_2 - rC$. A contract promising $(\overline{u}_1, \underline{u}_2)$ must ask player 2 to pay player 1, otherwise player 2's continuation payoff has to follow a diffusion process and violates his participation constraint with positive probability.[15] If $(\tilde{w}_1, \tilde{w}_2)$ are the continuation payoffs after payment, then

$$\overline{u}_1 + \underline{u}_2 = \tilde{w}_1 + \tilde{w}_2 - rC \leq \max_{(w_1,w_2)\in\mathcal{E}} w_1 + w_2 - rC.$$

The same argument at $(\underline{u}_1, \overline{u}_2)$ yields (2.14).

We can no longer follow the solution procedure in Section 2.2.6 (i.e., conjecture $\overline{u}_1$ and start the solution curve from $(\overline{u}_1, \underline{u}_2, \underline{\theta})$), because we do not know the initial angle $\underline{\theta}$.[16] Below, we propose a procedure that is suitable for the fixed-cost model.

We will need to conjecture two values: (1) the sum of payoffs $S := \max_{w_1} w_1 + f(w_1)$, and (2) the optimal w_1^* that satisfies the first-order condition

$$f'(w_1^*) = -1. \tag{2.15}$$

The advantage of this procedure is that the angle θ at w_1^* is $\frac{\pi}{4}$, thanks to (2.15). This allows us to solve the solution curve from $(w_1, w_2, \theta) = (w_1^*, S - w_1^*, \frac{\pi}{4})$ until it hits one of the boundaries **X**, **Y**.

For each pair (S, w_1^*), consider the solution to (2.10)-(2.12) with initial conditions $(w_1^*, S - w_1^*, \frac{\pi}{4})$. The solution curve will cross both **X** and **Y**,[17] and we denote the intersection points as $(\underline{u}_1, \overline{u}_2(w_1^*))$ and $(\overline{u}_1(w_1^*), \underline{u}_2)$, respectively.

Lemma 6. *For each* $S \in (\underline{u}_1 + \underline{u}_2, \max_{x\in[0,1]} p_1(x) + p_2(x))$*, there is a unique* $w_1^* \in$

[15] Because $(\overline{u}_1, \underline{u}_2)$ is an extreme point of the convex set $\mathcal{E}$, it cannot be delivered by any lottery.

[16] Alternatively, we can conjecture both $\overline{u}_1$ and $\underline{\theta}$ and search for a solution curve to satisfy (2.13) and (2.14). This procedure turns out to be numerically unstable because a solution curve never crosses the straight line **Y** (hence we cannot check equation (2.14)) whenever the conjectured $\overline{u}_1$ is too large or $\underline{\theta}$ is too small.

[17] This can be shown by an argument similar to that in Lemma 5.

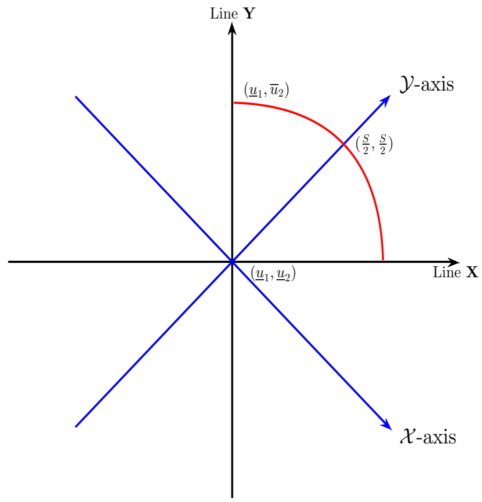

Figure 2.3: Change of coordinates.

$(\underline{u}_1, S - \underline{u}_2)$ *such that* $\underline{u}_1 + \overline{u}_2(w_1^*) = \overline{u}_1(w_1^*) + \underline{u}_2$.

The unique w_1^* in Lemma 6, which we denote as $w_1^*(S)$, allows us to reduce two value matching conditions (2.13)-(2.14) into one equation:

$$\underline{u}_1 + \overline{u}_2(w_1^*(S)) = S - rC. \tag{2.16}$$

How many solutions of S are there in (2.16)? Hinted by the linear-cost case, we conjecture two solutions. Although we have not been able to prove this result analytically, all of our numerical examples confirm this conjecture. Below, we offer a proof of this claim for a special case.

Suppose the two players are symmetric and $\underline{u}_1 = \underline{u}_2 = 0$. Because the Pareto frontier is symmetric around the 45-degree line, $w_1^* = w_2^* = S - w_1^* = S/2$. Given the symmetry, it is easier to work with new coordinates $(\mathcal{X}, \mathcal{Y})$

$$\mathcal{X} := \frac{\sqrt{2}}{2}(w_1 - w_2), \quad \mathcal{Y} := \frac{\sqrt{2}}{2}(w_1 + w_2).$$

That is, the $\mathcal{X}$-axis and $\mathcal{Y}$-axis under the new coordinates correspond to the negative-45-degree line and the 45-degree line under the old coordinates (see Figure 2.3). We parameterize the Pareto frontier as $\mathcal{Y} = -\frac{A}{2}\mathcal{X}^2 + \frac{\sqrt{2}}{2}S$, where $A > 0$ and $S > 0$ are parameters to be determined below. Two things are worth mentioning. First, the parameterized frontier is symmetric around the $\mathcal{Y}$-axis (i.e., the old 45-degree line). In particular, the peak of the frontier, $(0, \frac{\sqrt{2}}{2}S)$, corresponds to $(w_1^*, S - w_1^*) = (S/2, S/2)$ in the old coordinates. Second, for a given S, it is impossible for the frontier to satisfy the optimality equation at all points since there is only one free parameter, A. Therefore, we check the optimality equation only at $\mathcal{X} = 0$,[18]

$$\begin{aligned} A = \frac{-\mathcal{Y}''(0)}{\left(\sqrt{1 + (\mathcal{Y}'(0))^2}\right)^3} = \kappa(\pi/4) &= \max_{x \in [0,1]} \frac{2\mathbf{N}(\pi/4)((p_1(x), p_2(x)) - (S/2, S/2))}{r|\phi(\pi/4)|^2} \\ &= \frac{\sqrt{2}}{r|\phi(\pi/4)|^2}(2p(1/2) - S) \\ &= \sqrt{2}B(D - S), \end{aligned}$$

where $B := \frac{1}{r|\phi(\pi/4)|^2}$ and $D := 2p(1/2)$ are introduced to simplify notations. We derive (2.16) as follows. At $(\underline{u}_1 = 0, \overline{u}_2)$ in Figure 2.3, substituting $(\mathcal{X}, \mathcal{Y}) = (-\frac{\sqrt{2}}{2}\overline{u}_2, \frac{\sqrt{2}}{2}\overline{u}_2)$ into $\mathcal{Y} = -\frac{A}{2}\mathcal{X}^2 + \frac{\sqrt{2}}{2}S$, we have

$$\overline{u}_2 = \frac{-\frac{\sqrt{2}}{2} + \sqrt{\frac{1}{2} + 4\frac{A}{4}\frac{\sqrt{2}}{2}S}}{\frac{A}{2}} = \frac{\sqrt{1 + 2B(D-S)S} - 1}{B(D-S)},$$

[18] If one insists on solving the optimality equation everywhere, then there is no closed-form solution. Without a closed-form solution, it is difficult to check the number of solutions to (2.16).

where the second equality uses $A = \sqrt{2}B(D-S)$. Substituting $\underline{u}_1 = 0$ and the above into (2.16) yields

$$\frac{\sqrt{1+2B(D-S)S}-1}{B(D-S)} = S - rC. \tag{2.17}$$

Theorem 2. *Equation* (2.17) *has two solutions of* S *in* $(0, D)$.

Our analysis so far determines the Pareto frontier, but not the initial continuation values (W_{10}, W_{20}) to start the repeated game. Which point on the Pareto frontier is chosen usually depends on the players' bargaining powers. Here, however, we suppose that the game starts at $(w_1^*, f(w_1^*))$, i.e., the continuation values after a payment is made in our model. We can do this because the U.S. paid Canada in 1999.

2.3.2 Calibration

In this section, we will quantitatively evaluate the optimal contract. We first calibrate the model parameters to match certain observed features of the Pacific Salmon Treaty. We then compute the gain of switching to the optimal mechanism. Players 1 and 2 are, respectively, the U.S. and Canada. Throughout this section, our model period is one year (so the interest rate r is 0.04 in Subsection 2.3.2.3) and our results such as the welfare gains are measured in 1999 USD.

2.3.2.1 Estimation of μ, $p_1(\cdot)$, $p_2(\cdot)$

According to [18], illegal, unreported, and unregulated fishing may damage the environment and keep fisheries locked in low-value states. They estimate the environmental cost to be almost 9 billion euros when EU member states lose catches worth 10.7 billion euros. So we choose the value $\mu = \frac{9}{10.7} = 0.84$.

Next, we estimate Canada's payoff function $p_2(\cdot)$. Since our annual time series data are relatively short, we use monthly data to make our regression more accurate. That is, we first estimate the monthly payoff functions, and then we convert monthly payoffs into an annual payoff. This procedure consists of three steps.

1. Construction of monthly x_{2t} and p_{2t}. In the model, x represents the amount of the resource and p represents a player's payoff. In the data, we interpret x as the catching weight (in pounds) and p as the value added (in 1999 USD), which is defined as gross revenue minus the cost of intermediate goods such as repairs, gear, food, fuel, etc. The monthly catching weight for Canada is from Canada's Department of Fisheries and Oceans (DFO). The monthly value added for Canada is calculated as 22.5% of the monthly revenue reported by DFO. Appendix B explains how the value-added ratio is determined.

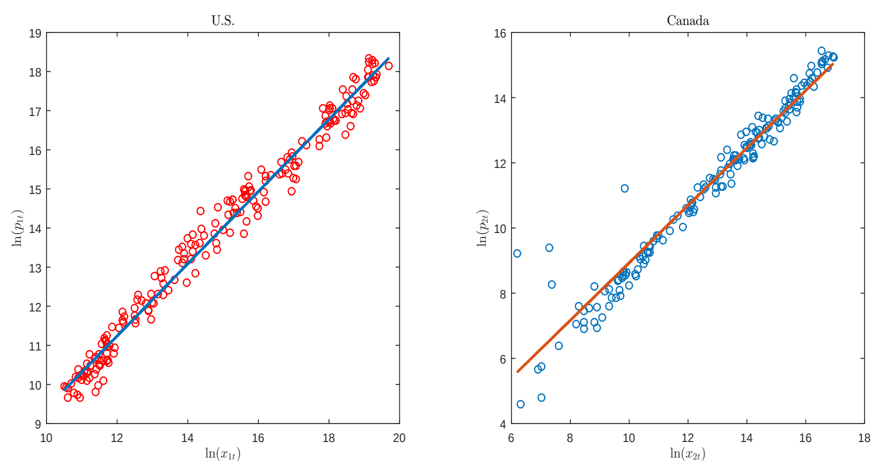

Figure 2.4: Approximately linear relationship between $\ln(p)$ and $\ln(x)$.

2. Estimation of monthly payoff function. The right panel in Figure 2.4 plots $\ln(p_{2t})$ against $\ln(x_{2t})$. It shows that $\ln(p_{2t})$ is approximately linear in $\ln(x_{2t})$. Therefore,

we run the following OLS regression[19]

$$\ln(p_{2t}) = \beta_2 + \alpha_2 \ln(x_{2t}) + \epsilon_{2t},$$

and assume that the monthly payoff function is $p_2^m(x_{2t}) = e^{\beta_2}(x_{2t})^{\alpha_2}$. Our regression results are in the last row of Table 2.1.

Table 2.1: Summary of OLS regression

Country	$\hat{\beta}_i$	$\hat{\alpha}_i$	Number of obs
U.S. ($i = 1$)	0.61	0.90	204
	(0.12)	(0.01)	
Canada ($i = 2$)	0.05	0.90	175
	(0.19)	(0.01)	

Notes: The U.S. data are provided by National Oceanic and Atmospheric Administration from 1999 to 2015. Canada's data are provided by its Department of Fisheries and Oceans from 1999 to 2014. Data with zero values are removed from regression.

3. Construction of annual payoff function. We define the annual payoff function as $p_2(x_2) := \sum_{k=1}^{12} p_2^m(\rho_k x_2)$, where $p_2^m(\rho_k x_2)$ is the payoff in the kth month, and ρ_k is the average ratio between the catching weight in the kth month and that in the whole year.

We follow the same procedure when we estimate $p_1(\cdot)$ for the U.S. In doing so, we face a missing-data problem that Alaska did not report its monthly data after 1998. We solve the problem as follows. Since the distributions of catching weight and revenue across different months were stable in Alaska before 1998, we assume that these distributions remain

[19]We also tried more general functional forms. For example, we added a quadratic term $(\ln(x_{2t}))^2$ as regressor, but found that the results in the following are barely changed.

unchanged after 1998. We then impute Alaska's monthly data after 1998 using Alaska's annual data and the unchanged distributions. Then, we construct the U.S. monthly data by aggregating the monthly data of Alaska and those of other states. Finally, we calculate the U.S. monthly value added as 40.9% of the monthly revenue, run the same linear regression for the U.S., and report our results in the first row of Table 2.1. Again, Appendix B explains how we determine the U.S. value-added ratio. The fact that the U.S. has a higher value-added ratio suggests that the U.S. production function is more efficient than Canada's, and this intuition will help us understand the welfare-gain result in Subsection 2.3.2.3.

2.3.2.2 Calibration of $(\sigma_1, \sigma_2, \underline{u}_1, \underline{u}_2, C)$ *and the total amount of resource* $\bar{x}$

This subsection calibrates the other six parameters $(\sigma_1, \sigma_2, \underline{u}_1, \underline{u}_2, C, \bar{x})$. Here $\bar{x}$ represents the total amount of resource; in the discussion so far, this amount has been normalized to 1.

Ideally, we want to build an extensive-form bargaining model to describe the formation of the Pacific Salmon Treaty (PST). Then we can calibrate the above six parameters so that this descriptive model replicates certain attributes of the treaty (payment amount, frequency, etc.). However, such a model is not immediately available because the bargaining process when the PST was signed in 1999 was not revealed to the public. Without much information about the actual bargaining process, we find it difficult to discipline any extensive form, and therefore think the following reduced-form model should serve our purpose equally well.

Suppose the two countries achieve some payoff pair $w \in \mathcal{E}$ after signing the PST. We allow w to be below the Pareto frontier because the PST may not be designed optimally due to contracting inefficiencies that we do not observe. To capture such inefficiencies, we assume that both countries have discount rate $\tilde{r}$ greater than r.[20] Under a higher discount

[20]The assumption that government is less patient than the market interest rate is also used in [19]. Political economy models (e.g., [20]) provide a justification for this assumption: incumbent politicians behave

Table 2.2: Calibrated parameters

σ_1	σ_2	$\underline{u}_1$	$\underline{u}_2$	C	$\bar{x}$	$\tilde{r}$
42.61	2.54	116.51	5.57	14.93	416.27	0.16

Notes: $(\sigma_1, \sigma_2, \underline{u}_1, \underline{u}_2, C)$ are measured in 1999 million USD. $\bar{x}$ is in million pounds.

rate, players are less patient and less willing to cooperate. On the other hand, we continue to assume that countries behave optimally in other aspects (i.e., sign an optimal contract under $\tilde{r}$). This way of modeling inefficiency is very parsimonious because all inefficiencies are embedded in one parameter $\tilde{r}$. We can adjust $\tilde{r}$ so that $w \notin \partial\mathcal{E}(r)$ lies on the Pareto frontier of a model with discount rate $\tilde{r}$, because the Pareto frontier and $\mathcal{E}(\tilde{r})$ shrink in $\tilde{r}$.

More specifically, we calibrate six parameters $(\sigma_1, \sigma_2, \underline{u}_1, C, \bar{x}, \tilde{r})$ to match six targets: (1) payment amount of the U.S.; (2) payment frequency of the U.S.; (3) mean of the U.S. payoffs; (4) variance of the U.S. payoffs; (5) mean of Canada's payoffs; (6) variance of Canada's payoffs. Note that we do not calibrate $\underline{u}_2$, because $\underline{u}_2$ affects mostly Canada's payment amount and frequency, but we have not observed any payment by Canada in our data set. We simply let $\underline{u}_2$ satisfy $\frac{\underline{u}_2}{\text{mean of Canada's payoffs}} = \frac{\underline{u}_1}{\text{mean of the U.S. payoffs}}$, i.e., Canada's outside option is similar to that of the U.S. (relative to their respective equilibrium payoffs).

The calibration results in Table 2.2 are in line with other related estimates. First, $(\sigma_1, \sigma_2) = (42.61, 2.54)$ is below the standard deviation of the U.S. payoffs and Canada payoffs because payoff variation contains both random shocks $\sigma_i dZ_{it}$ and the time varying $p_i(x_t)$. That $\sigma_1 \gg \sigma_2$ is not surprising because in the data both the standard deviation and the mean of the U.S. payoffs are much larger than those of Canada.[21] Second, the outside options $(\underline{u}_1, \underline{u}_2) = (116.51, 5.57)$ correspond to the case that countries receive

impatiently because they may not remain in power in the future.

[21] But the U.S. coefficient of variation is smaller than Canada's.

90% of their average annual income permanently. That is, compared with cooperative management, taking outside options and being noncooperative reduces countries' income by 10%. In the literature, estimates of the loss due to noncooperation range from 3% to 15%.[22] Our estimate is well within this range. Third, C is roughly equal to 5.46 years of operation cost of the Pacific Salmon Commission. Since the majority of the work done by the commission is to collect data on the salmon species and organize bilateral negotiation meetings, this estimate of C seems reasonable. Fourth, our estimate of $\bar{x}$ is equal to 95.20% of the two countries' average catching weight over 1999-2014. The discrepancy here might be due to the small-sample nature of the latter. Finally, our estimate of $\tilde{r}$ is significantly higher than r, suggesting that the underlying contracting inefficiencies are far from negligible.[23]

2.3.2.3 The optimal contract and its welfare gain

In this subsection, we compute the optimal contract and compare it with the positive contract in Subsection 2.3.2.2. The two contracts share the same parameter values in Table 2.2 except for the discount rate: $r = 0.04$ in the optimal contract,[24] while $\tilde{r} = 0.16$ in the positive one.

First, we compare the U.S. payment amount and frequency in the two contracts. Table 2.3 shows that the U.S. pays more frequently in the positive contract than in the optimal one. Since each payment incurs C, by paying less frequently the optimal contract incurs less total fixed cost, making it more efficient than the positive contract. One might wonder why the positive contract does not do the same (i.e., have a lower payment frequency and a larger payment amount). This is because under a high discount rate $\tilde{r} = 0.16$, the present

[22] See, e.g., [21], [22], [23], and [24].

[23] If one thinks that our estimate of $\tilde{r}$ is too high to be reasonable, please note that $\tilde{r}$ captures all sources of inefficiencies. $\tilde{r} = 0.16$ is chosen to match observables such as the high payment frequency of the U.S.

[24] We choose $r = 0.04$ because the implied annual discount factor is 0.96. This discount factor is commonly used in macroeconomics research.

Table 2.3: The U.S. payment amount and frequency

	Positive contract	Optimal contract
U.S. payment amount	\$56.31 million	\$327.58 million
Average duration between two U.S. payments	10 years	30.68 years

value of paying \$327.58 million every 30.68 years is much lower than paying \$56.31 million every 10 years, and thus the former payment scheme cannot compensate Canada sufficiently.

Second, we compare welfare levels achieved by the two contracts. For the optimal contract, instead of reporting total welfare of the two countries as $w_1 + w_2$, Table 2.4 reports $\frac{w_1+w_2}{r}$. Since $w_1 + w_2 = E\left[\int_0^\infty re^{-rt}\left(p_{1t}+p_{2t}\right)dt - \sum_{k=1}^\infty re^{-rt_k}C\right]$, where t_k is the kth payment time, by removing r in the flows and reporting $\frac{w_1+w_2}{r}$, we normalize our welfare level to be the standard present value of flows. Furthermore, to make this comparison meaningful, we calculate welfare in the positive contract using the same discount rate as in the optimal contract. That is, welfare in the positive contract is $E\left[\int_0^\infty e^{-rt}\left(p_{1t}+p_{2t}\right)dt - \sum_{k=1}^\infty e^{-rt_k}C\right]$, although $\{(p_{1t}, p_{2t}); t \geq 0\}$ and $\{t_k; k = 1, 2, 3, ...\}$ in the positive contract are derived under the discount rate $\tilde{r}$.

The first row in Table 2.4 shows that switching to the optimal contract improves welfare by \$51.66 million, or about 1.54%. There are two sources of this welfare gain: (1) the reduction in fixed cost, and (2) the increase in payoff flows. As discussed before, the optimal contract reduces the total fixed cost by paying less frequently. The second row in Table 2.4 confirms this: total fixed cost is reduced by \$23.79 millon, contributing to 0.71 percentage points in the welfare gain. More importantly, the third row in Table 2.4 shows that the present value of $p_{1t}+p_{2t}$ is higher by \$27.87 millon in the optimal contract, which

Table 2.4: Change of welfare measured in 1999 million USD

	Positive contract	Optimal contract	Difference
$E\left[\int_0^\infty e^{-rt}(p_{1t}+p_{2t})\,dt - \sum_{k=1}^\infty e^{-rt_k}C\right]$	3351.36	3403.02	51.66
$E\left[\sum_{k=1}^\infty e^{-rt_k}C\right]$	29.58	5.78	-23.79
$E\left[\int_0^\infty e^{-rt}(p_{1t}+p_{2t})\,dt\right]$	3380.94	3408.81	27.87
$E\left[\int_0^\infty e^{-rt}p_{1t}dt\right]$	3226.06	3385.9	159.84
$E\left[\int_0^\infty e^{-rt}p_{2t}dt\right]$	154.88	22.91	-131.97

contributes to 0.83 percentage points in the welfare gain. Since the increase in $p_{1t} + p_{2t}$ explains more than half of the welfare gain, in the following we shall try to understand the reason for this increase.

The last two rows of Table 2.4 report separately the present values of p_{1t} for the U.S. and p_{2t} for Canada. The U.S. receives more payoff flows in the optimal contract than in the positive contract, while Canada does the opposite. This implies that the optimal contract reallocates salmon resources from Canada to the U.S. In fact, the average U.S. catching weight has increased by 21.4 million pounds moving from the positive contract to the optimal contract, while Canada's weight has decreased by the same amount. Note that the U.S. gains more in terms of value added than Canada loses, so the total payoff increases after this reallocation. That the U.S. production function is more efficient than Canada's is consistent with the earlier observation that the ratio between value added and revenue is 40.9% in the U.S. but only 22.5% in Canada.

2.4 Conclusion

In this paper, we study the optimal design of the Pacific Salmon Treaty under two-sided moral hazard. We extend the theory of continuous-time repeated games to allow for side payments. We view our analysis as making two contributions. First, we show that there are only two solutions to the optimality equation that characterizes the boundary of the set of equilibrium payoffs. This technical contribution greatly simplifies the calculation of the equilibrium set. In the literature, the search for the solution to the optimality equation is done by trial and error. Second, we provide a useful policy recommendation to improve upon the Pacific Salmon Treaty. Because the U.S. production function is more efficient, our optimal contract would give a bigger salmon share to the U.S. than the current treaty does. This policy change will improve the two countries' welfare by 1.54%.

It is possible to extend our model and add the stock of the salmon population to our repeated game as a new state variable. Naturally, the optimal sharing rule should depend on this variable. We conjecture that our main finding (that the U.S. ought to get a bigger share of the resource due to production efficiency) is robust to this extension.

We can also extend the current contracting framework to include three or more countries. This extension is useful in practice because many government negotiations are multilateral; for example, Western Pacific salmon are shared by China, Japan, Russia, and South Korea. However, the challenge here lies in theory: the optimality equation with three or more players is a partial differential equation and finding its boundary condition may not be easy.[25] We leave this extension to future research.

[25] In a multi-agent moral-hazard model, [25] reduces the optimality equation to an ordinary differential equation by assuming that all agents are symmetric, i.e., they exert the same effort and receive the same consumption. This assumption, however, is too restrictive for our application: the U.S. salmon industry is more than 10 times bigger than Canada's.

3. BOOM AND BUST IN CHINA'S HOUSING MARKET

3.1 Introduction

China's housing price has been growing steadily over the past decade. The average growth rate of housing price is 10% from 2003 to 2013, far exceeding the 1.4% growth rate of rents. The high price/rental ratio suggests that the current housing price cannot be fully explained by the discounted sum of future rents, i.e., there is a bubble in the housing market. Moreover, the rate of return from housing asset is quite different from the return from capital. While the growth rate of housing price remains steadily at 10%, the return to capital has been declining, reaching a low level of 5% in 2013. In an equilibrium, investors holding both housing and capital must be indifferent between the two options. No-arbitrage condition therefore implies that the high growth rate of housing price is also coupled with high risk of price crash. This risk of bubble burst has drawn a lot of attention from policy makers, social media, and academia.

Our paper studies bubble burst. Following [26], we model China's housing market as one of rational bubble in an overlapping generation framework. Young entrepreneurs use their endowment to either purchase housing or invest in firm's capital. The economy is initially in a bubbly state where housing price is above the fundamental value. In each period, a sunspot equilibrium decides whether the bubble continues to grow, or the bubble bursts and housing price falls back to the fundamental value. In the latter case, less expensive housing allows young entrepreneurs to allocate more resource to capital investment, dampening the effect that housing bubble crowds out private capital.

Our model features another channel through which housing price may affect the economy. Data from 2003 to 2013 show that around 45% of China's infrastructure investment is funded by the government's revenue from land auction to the private sector. This motivates us to model a government who chooses the level of infrastructure investment based on its revenue from land sale. Clearly, a bubble burst will reduce land price and the government's infrastructure investment. What is less clear is the effect of lowered infrastructure on China's GDP. If the infrastructure investment is already excessive, then bubble burst will improve production efficiency by reducing the excessiveness of infrastructure. On the other hand, if infrastructure is inadequate, then bubble burst will make the situation even worse.

To quantify the effects of housing price and infrastructure on the economy, we calibrate our model to match growth rates of both the real GDP and housing price of China. Two findings from our calibrated model are worth mentioning. First, we confirm the existence of bubble in China's housing market. Since housing price has been growing faster than rents for a decade, bubble size has gone up from 3% of the housing price in 2003 to 32% in 2013. If this bubble were to crash in 2017, housing price will drop by 40% and entrepreneurs' total wealth drop by 16%. Second, we find that China's infrastructure is indeed overinvested. While the optimal ratio between infrastructure and private capital is 1:4.5, this ratio in China is 1:3.7 based on our estimates.[1]

We use the model to answer two important questions: 1) what is the consequence of a housing market crash? and 2) how does the adoption of property tax affect housing market? In the first question, we suppose a shock that eliminates the housing bubble occurs in 2017, and then simulate the equilibrium dynamics afterwards. Unsurprisingly, market price of all existing homes takes a big hit. Since newly built homes enter GDP, China's GDP growth rate decreases from 6% to 2.3% in 2017. This decline, however, is not long lasting: China's GDP after the crash of 2017 would overtake what would have been achieved if the shock had not occurred by 2047. The reason for the quick

[1] China does not report the stocks of infrastructure or private capital, so a model must be used to estimate them.

recovery is that lowered housing price reduces the government's revenue from land sale, and consequently overinvestment of infrastructure can be resolved. With more capital invested in non-housing sectors, higher output from these sectors makes up for the lost value of new homes. In fact, China's GDP excluding housing is unaffected by the crash of 2017, and by 2047 is higher than that with no crash by 5%.

To answer the second question, we suppose that Chinese government imposes a 1.5% property tax on all homes starting from 2017. Even if the bubble does not burst, this policy will immediately reduce the bubble size in 2017 because the after-tax return of owing a home is lower. In fact, the rent-to-price ratio goes from 1.3% to 3.2%. More importantly, we find that output of non-housing sectors increases more than it does under the crash of housing bubble. The reason is that government's revenue (and its infrastructure investment) are affected differently in the two cases. In the case of bubble burst, the reduction of revenue is so dramatic that infrastructure goes from overinvestment to underinvestment. In the case of property tax, however, the reduction of government revenue is less dramatic because the property tax revenue partially offset the loss in land-sale revenue.

Related literature Our paper is closely related to [27]. Both papers treat Chinese housing as an asset bubble and analyze the movement of housing price in rational bubble framework. There are, however, three differences between our paper and [27]. First, the only effect of housing bubble in [27] is to crowd out the investment on productive capital and slow down capital accumulation. In addition to that effect, our paper also studies the infrastructure effect that housing bubble increases the government's land-auction income and enhances its infrastructure investment. Second, in [27], capital return is constant during the first stage of the economy. In our model, capital return can vary overtime, which is more consistent with the fact that Chinese capital return dropped in the last decade. Third, our model assumes a nonzero probability of bubble burst in each period. This assumption helps us match the housing price dynamics documented in [28].

There is a large literature studying how bubble burst affects the economy, including [29], [30], [31], and [32]. A common theme of these papers is the focus on firms' credit constraints. In these papers, housing assets are used by firms as collateral and a bubble burst tightens the firms' credit constraints, forcing them to take inefficiently low investment. Although credit constraint is important in understanding bubble burst in the 1989 Japanese housing market and the 2007 U.S. housing market, it is less so in China for two reasons. First, the chance that bubble burst triggers a banking crisis is low because mortgage loans in China require high down payments ([28]). Second, large state owned enterprises do not rely on housing to get loans while most private enterprises are excluded from the financial market (see [33]).

Our paper is also related to a large literature on government's expenditure and economic growth, e.g., [34], [35], [36], [37], [38]. As in the literature, efficiency in our model requires a good balance between infrastructure investment and private capital investment. Our model differs from the literature in modelling the government's budget constraint. While infrastructure is purely funded by tax revenue in the literature, here it is also funded by government's sale of a bubbly asset. We emphasize this channel because almost one fourth of Chinese government's income comes from land sale.

Our paper is organized as follows. Section 3.2 describes the basic model, and characterize its equilibrium with housing bubbles. Section 3.3 extends the basic model with population and technology growth. The extended model is calibrated in Section 3.4, and used in Section 3.5 to study the consequence of a bubble burst and the adoption of property tax. Section 3.6 concludes. The Appendix contains proofs of all the results.

3.2 Basic Model

We follow [26] to build an overlapping-generation model. In this basic model, there are no population growth or technology improvement. Moreover, we assume housing is a pure bubble to simplify our analysis. All these assumptions are relaxed in the next section, where we do calibration.

3.2.1 Environment

There are three types of agents in this model: workers, entrepreneurs and a government. Workers and entrepreneurs live for two periods. Each period, a unit measure of new workers and entrepreneurs enter the model to replace the old. Young workers have one unit of inelastic labor. After receiving wage income, workers decide their consumptions and savings. They are out of capital market and housing market. The only way to save is to purchase risky-free government bond. Their optimization problem is

$$\begin{aligned} \max \quad & \log(c^w_{1t}) + \rho \log(c^w_{2t}) \\ s.t. \quad & c^w_{1t} + b_t = w_t, \\ & c^w_{2t} = R^f b_t, \end{aligned}$$

where c^w_{1t} and c^w_{2t} are consumptions for workers born in period t, ρ is the discount factor, w_t is wage, b_t is the holding of risk-free bond, and R^f is the exogenous interest rate set by the government.

To understand entrepreneurs' problem, we first describe the dynamics of housing price Q_t. In this section, housing does not generate rents and therefore is a pure bubble. People buy it only because they can resell it in the future. The economy is either in a bubbleless state ($Q_t = 0$), or in a bubbly state ($Q_t > 0$). If $Q_t = 0$, then $Q_s = 0$ for all $s \geq t$ (i.e., the bubbleless state is absorbing). If $Q_t > 0$, then the bubble bursts with probability $1 - p$ in period $t+1$:

$$Q_{t+1} = \begin{cases} Q^b_{t+1} > 0, & \text{with prob } p; \\ 0, & \text{with prob } 1-p. \end{cases}$$

How Q^b_{t+1} depends on Q_t is endogenous and will be studied below.

Entrepreneurs can invest in both capital and housing. They are risk neutral and only care about their second-period consumption. Each young entrepreneur inherits an initial wealth of m_t from the older generation. His problem is

$$\begin{aligned} \max \quad & E[c^e_{2t}] \\ s.t. \quad & k_{t+1} + Q_t h_t = m_t, \\ & c^e_{2t} = R_{t+1} k_{t+1} + Q_{t+1}(1-\delta) h_t, \end{aligned}$$

where h_t is the amount of house a young entrepreneur purchases, k_{t+1} is the investment in his firm, R_{t+1} is the capital return rate at time $t+1$, and δ is the depreciation rate of house.

Each old entrepreneur owns a firm after investment of k_t. The firm's production function depends on the aggregate infrastructure level, A_t, as follows

$$y_t = \hat{A}_t^\theta k_t^\alpha l_t^{1-\alpha},$$

where k_t and l_t are each firm's capital and labor. $\hat{A}_t := \frac{A_t}{K_t^\beta L_t^{1-\beta}}$ is the aggregate infrastructure level adjusted for congestion effect. red We follow [37] to build this production function. As [37] points out, in many cases, such as highways, utilities, and bridges, productivity of infrastructure would decrease when more people or firms use them. Thus we assume the productivity of infrastructure should be adjusted by the aggregate capital K_t and aggregate labor L_t.

As in [37], we assume

Assumption 2. $\alpha - \beta\theta > 0$ *and* $\alpha + (1-\beta)\theta \leq 1$.

The first inequality guarantees the marginal return of capital is positive. The second inequality guarantees the return on the whole reproductive part (infrastructure and capital) is weakly decreasing. After production, old entrepreneurs pay tax to government at rate τ and a fixed fraction ψ of post-tax production to young entrepreneurs as their initial wealth, i.e., $m_t = \psi(1-\tau)y_t$. The remainder is dividend, which old entrepreneurs consume. The old entrepreneur's optimization problem is

$$\begin{aligned} D_t &= \max_{l_t}[(1-\tau)(1-\psi)\hat{A}_t^{\theta} k_t^{\alpha}(l_t)^{1-\alpha} - w_t l_t], \\ &= R_t k_t, \end{aligned}$$

where D_t is dividend entrepreneur gets at time t, and R_t is capital return.

Capital in the economy evolves as

$$k_{t+1} = (1-\phi_t)m_t = (1-\phi_t)(1-\tau)\psi y_t, \tag{3.1}$$

where $\phi_t := \frac{Q_t h_t}{k_{t+1}+Q_t h_t}$ denotes the fraction of housing in young entrepreneurs' portfolio.

A government supplies ΔH_t units of housing to the market exogenously, and also invests in infrastructure. The government budget constraint is

$$A_{t+1} = (\tau - e)\hat{A}_t^{\theta} K_t^{\alpha} + Q_t \Delta H_t, \tag{3.2}$$

where e is the ratio of government expenditure on some public good other than infrastructure. We assume full depreciation of both capital and infrastructure in our basic model for convenience. All the properties shown will continue to hold when the depreciation rate is smaller than 1. The total amount of houses H_t evolves as

$$H_{t+1} = \Delta H_t + (1-\delta)H_t. \tag{3.3}$$

To guarantee the steady state and the sustainable bubble exist, we assume

$$\lim_{t\to\infty} g_t = 1,$$

where $g_t := \frac{H_{t+1}}{H_t}$.

3.2.2 Equilibrium

This subsection characterizes equilibrium of the economy. We start with its definition.

Definition 1. *An equilibrium is a sequence of consumptions* $\{c_{1t}^w, c_{2t}^w, c_t^2\}_{t=0}^\infty$, *savings* $\{b_t, k_t, h_t\}_{t=0}^\infty$, *labor supply/demand* $\{l_t\}_{t=0}^\infty$, *infrastructure* $\{A_t\}_{t=0}^\infty$ *and prices* $\{w_t, R_t, Q_t\}_{t=0}^\infty$ *such that 1) workers and entrepreneurs maximize life-time utilities, 2) firms maximize profits, 3) government's budget constraint is satisfied, and 4) the labor, capital, and housing markets clear. In particular, we have*

- $l_t = 1, \forall t$ *because labor supply is inelastic;*
- *first-order conditions for firms' profit maximization problem (after imposing* $l_t = 1$*):*

$$w_t = (1-\alpha)(1-\tau)(1-\psi)\frac{Y_t}{L_t} = (1-\alpha)(1-\tau)(1-\psi)\hat{A}_t^\theta K_t^\alpha, \tag{3.4}$$

$$R_t = \alpha(1-\tau)(1-\psi)\frac{Y_t}{K_t} = \alpha(1-\tau)(1-\psi)\hat{A}_t^\theta K_t^{\alpha-1}; \tag{3.5}$$

- *non-arbitrage condition for young entrepreneurs to invest in both housing and cap-*

ital:

$$\frac{pQ_{t+1}^b(1-\delta)}{Q_t^b} = R_{t+1}. \tag{3.6}$$

The main characterization result of this section is as follows.

Proposition 1. *Consider an economy with initial condition $\{k_0, A_0, H_0\}$. If $\frac{(1-\psi)\alpha}{(1-\delta)\psi} \geq p$, then no bubbly equilibrium exists, i.e., $Q_t = 0$ for all dates and states in the equilibrium. Otherwise, if $\frac{(1-\psi)\alpha}{(1-\delta)\psi} < p$, then a continuum of equilibria exist depending on the initial Q_0. There is a $\hat{Q}_0^b > 0$ such that*

1. *if $0 < Q_0 < \hat{Q}_0^b$, then a bubbly equilibrium exists in which $\lim_{t\to\infty} Q_t^b = 0$;*
2. *if $Q_0 = \hat{Q}_0^b$, then a bubbly equilibrium exists in which $\lim_{t\to\infty} Q_t^b > 0$;*
3. *if $Q_0 > \hat{Q}_0^b$, then no equilibrium exists.*

The intuition for Proposition 1 can be explained in two steps. First, we explain how the long-run size of the bubble, $\lim_{t\to\infty} Q_t^b$, depends on the initial $Q_0^b = Q_0$. With higher Q_0^b, more private capital K_1 is crowded out, and capital return R_1 becomes higher. Non-arbitrage condition (3.6) then implies a higher growth rate $\frac{Q_1^b}{Q_0^b}$. Using this argument for all the future dates, we conclude that higher Q_0 increases the growth rate $\frac{Q_{t+1}^b}{Q_t^b}$ for all t. Now, let $\hat{Q}_0^b$ be the initial bubble size starting from which the bubble stabilizes in the long run, i.e., $0 < \lim_{t\to\infty} Q_t^b < \infty$. Then in case (i), lower Q_0 and lower growth rates imply the bubble eventually disappears, while in (iii), higher Q_0 and higher growth rates imply the bubble eventually explodes. Note that in (iii) equilibrium does not exist with exploding housing price, because young entrepreneurs, whose initial wealth are bounded, eventually cannot afford to purchase the bubble (thus violating the market-clearing condition in housing).

Second, we discuss two senses in which condition $\frac{(1-\psi)\alpha}{(1-\delta)\psi p} < 1$ is needed for the existence of bubbly equilibrium. On the one hand, because the bubble size $\phi^* := \lim_{t\to\infty} \phi_t$ in a bubbly steady state equals $1 - \frac{(1-\psi)\alpha}{(1-\delta)\psi p}$, we need $1 - \frac{(1-\psi)\alpha}{(1-\delta)\psi p} > 0$ for bubbly steady state to exist.[2] On the other hand, because $\frac{(1-\psi)\alpha}{\psi}$ is the long-run real interest rate in the bubbleless economy,[3] the condition $\frac{(1-\psi)\alpha}{(1-\delta)\psi p} < 1$ is nothing more than an upper bound imposed on this real interest rate, which is a standard assumption in the literature for bubbles to exist. As pointed out by [26], only if the rate of return from capital (in the absence of bubble) is sufficiently low, bubbly assets may enter the economy as an alternative channel to save.

3.2.3 Two Steady States

There are two steady states in our model: one is bubbleless and the other is bubbly. In this subsection, we compare output levels of the two steady states.

In both steady states, output Y^*, infrastructure A^* and private capital K^* satisfy

$$
\begin{aligned}
Y^* &= (A^*)^\theta (K^*)^{\alpha-\beta\theta}, \\
A^* &= (\tau - e + \delta(1-\tau)\psi\phi^*)Y^*, \\
K^* &= \psi(1-\tau)(1-\phi^*)Y^*.
\end{aligned}
$$

Bubble size ϕ^* equals some positive value $\phi_b > 0$ in the bubbly steady state, but is zero in the bubbleless steady state. Clearly, large bubble size ϕ^* enhances A^* by crowding out private capital K^*.

[2] That $\phi^* = 1 - \frac{(1-\psi)\alpha}{(1-\delta)\psi p}$ can be derived from the following conditions:

$$
\begin{aligned}
K^* &= \psi(1-\tau)(1-\phi^*)Y^*, &(3.7)\\
R^* &= \alpha(1-\tau)(1-\psi)\frac{Y^*}{K^*}, &(3.8)\\
p(1-\delta) &= R^*, &(3.9)
\end{aligned}
$$

where (3.7), (3.8), and (3.9) are long-run limits of (3.1), (3.5), and (3.6).

[3] That $R^* = \frac{(1-\psi)\alpha}{\psi}$ in the bubbleless economy can be derived from (3.7), (3.8), and $\phi^* = 0$.

Output in the bubbly steady state is higher than that in the bubbleless steady state if and only if

$$\left(1+\frac{\delta(1-\tau)\psi}{\tau-e}\phi_b\right)^{\theta}(1-\phi_b)^{\alpha-\beta\theta}>1. \tag{3.10}$$

In (3.10), $\frac{\delta(1-\tau)\psi}{\tau-e}\phi_b$ captures the (percentage) increase in infrastructure funded by the sale of bubbly asset, while $-\phi_b$ captures the decrease in capital. To obtain further intuition, suppose ϕ_b is small, say 1%. Then the change of the first term, $\left(1+\frac{\delta(1-\tau)\psi}{\tau-e}\phi_b\right)^{\theta}$, from its bubbleless-state value of one is $\frac{\delta(1-\tau)\psi}{\tau-e}\theta$ percent, while the change of the second term, $(1-\phi_b)^{\alpha-\beta\theta}$, is $-(\alpha-\beta\theta)$ percent. Then (3.10) holds if and only if

$$\delta\frac{\psi(1-\tau)}{\tau-e}\frac{\theta}{\alpha-\beta\theta}>1.$$

The intuition for the above is as follows. First, $\frac{\psi(1-\tau)}{\tau-e}$ is the ratio $\frac{K^*}{A^*}$ in the bubbleless steady state, and a higher ratio increases the return of reallocating capital to infrastructure through bubble, making the bubbly output higher than the bubbleless output. Second, $\frac{\theta}{\alpha-\beta\theta}$ is the ratio between the elasticities of infrastructure and capital. If infrastructure is more elastic than capital, then again the return of reallocating capital to infrastructure is higher. Third, the depreciation rate of housing, δ, is the ratio between new housing ΔH and total housing H in any steady state. If one unit of capital is crowded out by bubble, only δ units enter the revenue of the government and become infrastructure investment (the rest, $1-\delta$, belongs to sellers of old housing units). That is, δ is the rate of transformation between K^* and A^*, and a higher δ increases the return of reallocating capital to infrastructure.

The above intuition continues to hold when ϕ_b is not small. In fact,

Corollary 2. *Inequality* (3.10) *is more likely to hold with higher* $\frac{\psi(1-\tau)}{\tau-e}$, $\frac{\theta}{\alpha-\beta\theta}$, *and* δ.

Figure 3.1 further illustrates the dependence of outputs on the government expenditure

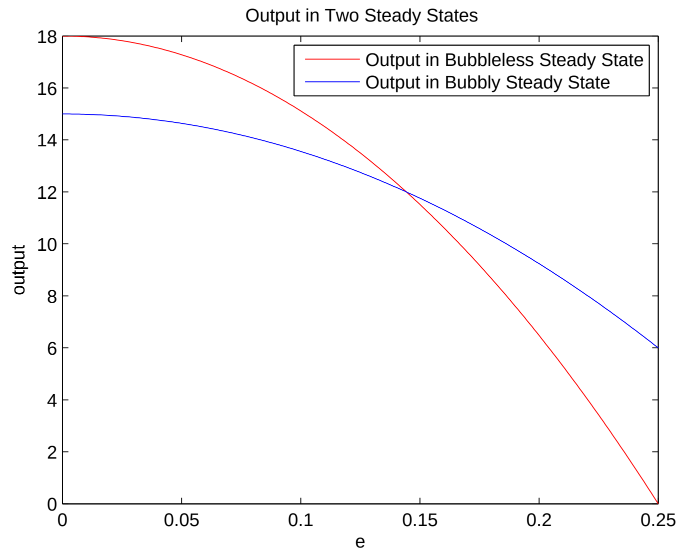

Figure 3.1: Output of bubbly steady state and bubbleless steady state.

e. When e is closed to zero, all the fiscal income is used for infrastructure investment and infrastructure is over-invested. The bubbly output is lower than the bubbleless output because bubble worsens the over-investment problem. When e is close to τ, however, the bubbly output is higher than the bubbleless output because the latter is close to zero when the government cannot provide any infrastructure.

In [26] and [27], bubble only has the effect of crowding out capital, which certainly lowers the output. In our paper, however, bubble also helps government accumulate infrastructure. Thus it is possible that the bubbly output is higher than the bubbleless output.

3.3 Extended Model

We build an extended model to better approach to the reality. There are four main differences comparing to the basic model. First, in this extended model, we consider an economy with population growth and technology improvement. The Population of both workers and entrepreneurs grows at a constant rate $1 + \eta$, and the labor efficiency grows at a constant rate $1+\zeta$. Second, both workers and entrepreneurs live for T years, and workers retire at the age J. Third, we consider house delivers real utility flow r_t, thus house has some fundamental value. The rent r_t grows at a exogenous growth rate $1+\xi$. Forth, capital and infrastructure depreciate at $1-\delta^k$ and $1-\delta^a$.

The j-th generation workers' optimal decision problem can be written as,

$$\max E[\sum_{t=j}^{T+j} \rho^{(t-j)} log(c^w(j,t))]$$

subjected to,

$$\begin{aligned} c^w(j,j) + b^w(j,j) &= w_j \\ c^w(j,t) + b^w(j,t) &= w_t + R^f s^w(j,t-1), \text{when} \in [j+1, j+J] \\ c^w_(j,t) + b^w(j,t) &= R^f b^w(j,t-1), \text{when } t \in [j+J+1, j+T], \end{aligned}$$

where $c^w(j,t)$ stands for j-th generation worker's consumption at time t and $b^w(j,t)$ stands for j-th generation worker's saving at the beginning of time t.

The j-th generation entrepreneurs are born with initial endowment m_j, and choose their portfolio between capital investment and housing investment. The entrepreneurs are risk-neutral and care about their consumption in last period. Their optimal decision problem can be written as.

$$\max E[c^e(j, j+T))]$$

subjected to,

$$\begin{aligned} Q_j h(j,j) + k(j,j+1) &= m(j), \\ Q_t h(j,t) + k(j,t+1) &= R_t k(j,t-1) + (Q_t + r_t) h(j,t-1)(1-\delta), t \in [j+1, j+T-1] \\ c^e(j, j+T)) &= R_t k(j, j+T-1) + (Q_t + r_t) h(j, j+T-1)(1-\delta). \end{aligned}$$

where $c^e(j,t)$ denote j-th generation entrepreneurs' consumption at time t. $h(j,t)$ and $k(j,t)$ stand for the quantity of housing and quantity of capital j-th generation entrepreneurs own at the beginning of time t.

A non-arbitrage condition in this extended model is

$$\frac{[pQ^b_{t+1} + (1-p)Q^f_{t+1} + r_t](1-\delta)}{Q^b_t} = R_{t+1},$$

where Q^f_t is the fundamental price of housing, which is defined as $Q^f_t \equiv \sum_{i=t+1}^{\infty} \frac{r_i(1-\delta)^{i-t}}{\prod_{j=t}^{i} R_j}$.

Old entrepreneurs face similar problem in the basic model. The only difference is that we consider depreciation rate of capital is δ_k and labor efficiency E_t in this extended model. The optimal problem faced by old entrepreneurs as

$$\max_{l_t} (1-\psi)\hat{A}^\theta_t K^\alpha_t (E_t l_t)^{1-\alpha} + (1-\delta_k) K_t.$$

The government's budget constraints is

$$A_{t+1} = (\tau - e)\hat{A}^\theta_t K^\alpha_t + \kappa Q_t \Delta H_t + (1-\delta_a) A_t,$$

where δ_a is the depreciation rate of infrastructure capital and κ is the fraction between net revenue of selling land and the total value of newly built house.

3.4 Calibration

3.4.1 Parameters

Our calibrated model focuses on matching Chinese data during 2003-2013. Each period in our model stands for one year in reality, and the model starts at 2003. There are two types of parameters needed to be decided. The first type of parameters are chosen exogenously and the other type are calibrated in the model.

We firstly introduce parameters chosen exogenously in this and next paragraphs. Similar with [33] and [27], agents enter into the economy at age 22 and live for $T = 50$ years, which is consistent with the average life expectancy 71.4 years from 2000 Chinese Population Census. Workers retires after 30 years working. The population growth rate is 0.005, which matches the average population growth rate during 2003-2013 from NBSC data set. $R^f =$ is set as 1.0175, matching the average one-year real deposit rate. ρ is chosen as 0.997 to match the average investment ratio.

On the production side, the capital income share α is chosen as 0.5, which is consistent with [39]. θ is chosen as 0.1, which is estimated by [40]. Since we find β is no sensitive in the model, thus we assume the congestion effect to capital and to population are the same and β is conjectured as 0.5. Following [39], the depreciation rate of capital δ_k and the depreciation rate of infrastructure δ_a are set as 0.1.

The growth rate of rent g_r is set as 0.014 to match the average growth rate of rent during 2003-2013. The growth rate of housing g_t is decided as following. In the first eleven years we set $g_t = 0.07$ to match the fact that quantity of house doubled during 2000-2010. After the first 11 years, we think it seems impossible that such high growth rate of housing can last for long term since Chinese growth rate of urban population has already decreased for last decade. We assume the long-run growth rate of housing equals to the growth rate of population, and the growth rate of housing follows an exponential attenuation function as $g_t = (g_0 - \eta)e^{-\xi(t-11)} + \eta$, where attenuation speed $\xi = 0.07$ to match the attenuation speed of the urban population growth during 2003-2013.

After we introduce parameters chosen exogenously, now we show parameters estimated in the model. ψ is chosen as 0.19 to match the capital return before tax 22 percent in 2003. τ is chosen as 0.21 to match the capital return after tax 15 percent in 2003. e is chosen as 0.16 to match the proportion of government expenditure (excluding infrastructure investment) out of GDP. The growth rate of labor efficiency ζ is set as 0.04 to match 10 percent Chinese average GDP growth rate during 2003-2013.

We use the possibility of bubble burst to match the growth rate of housing price. And we assume the possibility of bubble burst degenerates in long run to guarantee the existence of bubble. For simplicity, we assume that the probability of bubble burst is a linear attenuation function with time. The starting probability of bubble burst and ending time are set as $1 - p = 0.17$ and $T = 15$ to match the average growth rate of housing price during 2003-2008 and 2009-2013.

The initial labor quantity and housing quantity are normalized as 1. The initial aggregate capital level is set as 1.41 to match the capital-to-output ratio 1.26 in 2003. The initial infrastructure is set as 0.68 to match the ratio of infrastructure to capital 0.48 in 2003 (see [41]). Following [33] the initial wealth distribution of entrepreneurs is set as wealth distribution of workers in steady state. The initial housing rent is set as 0.017 to match the proportion of newly built house value to GDP 4.5% in 2003.

3.4.2 Calibration Result

Our main calibration result is shown in Figure 3.2. In the left panel, we replicate the path of capital return. Our simulated path is consistent with capital return data in [42]. Based on [42], capital return in China drops from 15 percent in 2003 to 5 percent in 2013.

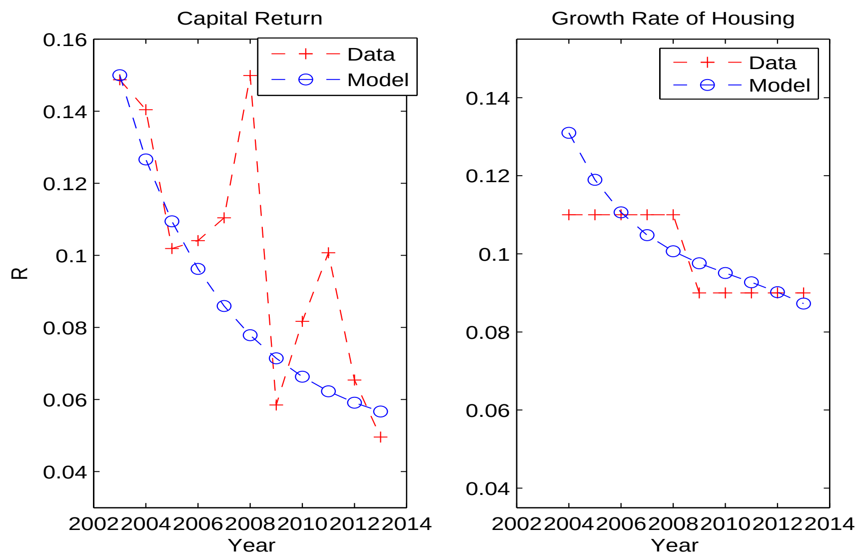

Figure 3.2: Return to capital and housing price in China.

Our simulated capital return drops from 15 percent to 5.8 percent, which reasonably well matches the rapid decrease on capital return since 2003. Our simulated growth rate of housing pricing drops from 13 percent in 2003 to 8.7 percent in 2013. The Decrease on growth rate of housing price is only one half of the decrease on capital return, which matches the growth rate of housing in [28]. The resilient growth rate of housing price is the key prediction in our model.

In the Figure 3.3, we show GDP growth rate and infrastructure investment in our model and in the data. The left panel of Figure 3.3 shows GDP growth rate. Our model can replicate the average GDP growth rate as 10 percent during 2003-2013, and also the path that the growth rate drops from the highest point over 13 percent to 7.5 percent in 2013. In the right panel of Figure 3.3, we show the increasing proportion of infrastructure investment to GDP. Both our model and the data suggest that this proportion increased dramatically: in the data, it increased from 7 percent to 12.6 percent; in our model, it increased from 8.7 percent in 2003 to 11 percent in 2013. Our model may underestimate the high proportion

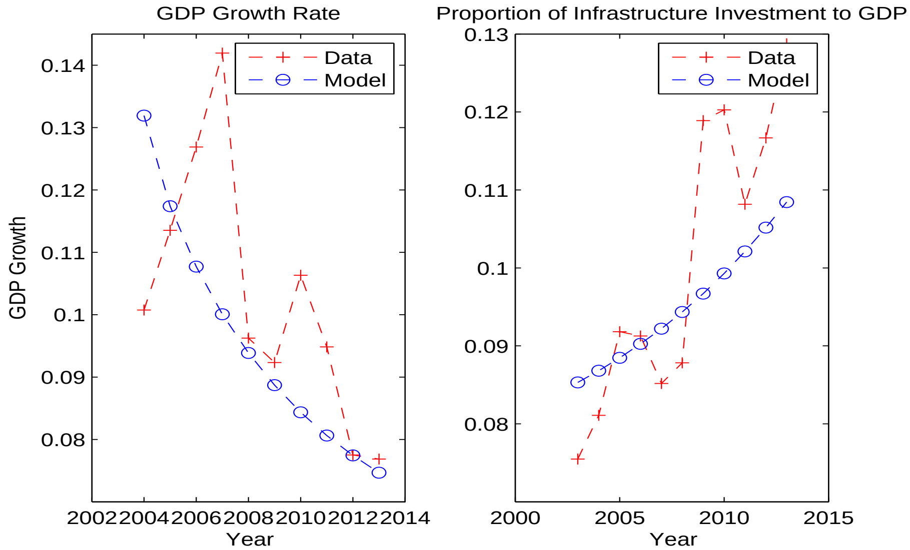

Figure 3.3: GDP growth and infrastructure investment.

of infrastructures investment because after financial crisis Chinese government proposed a large economic stimulus package (which is well known as "4 trillion project"). Much fund from this project was invested into infrastructure building. This can explain the underestimation after 2008 in our model.

In the Figure 3.4, our model predicts the proportion of bubble component to total housing price. Through our simulation, the proportion of bubble component increases dramatically from 5 percent in 2003 to 32 percent in 2013. This dramatic increase comes from the fact that the average growth rate of housing price is around 10 percent while the average growth rate of rent is only 1.4 percent. The growth of bubble proportion is the main reason why our model can predict resilient growth of housing price: when the bubble component is small, even if public predict the possibility of bubble burst, the growth rate of housing price is not affected by this prediction and lower than capital return because total housing return includes both growing value and rent; when the bubble component is large, people's expectation on the possibility of bubble burst becomes dominating power

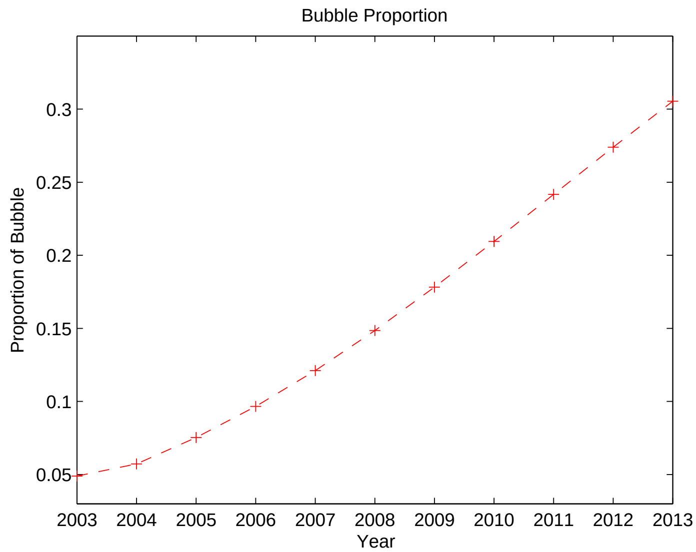

Figure 3.4: Bubble proportion.

to push the growth of bubble even higher to compensate for the risk people take.

3.5 Counterfactual Experiment

3.5.1 Crowding-out effect and Infrastructure Effect

To better understand how bubble affects the economy, we conduct a counterfactual experiment to explore the two effects in our model. We assume an economy which is almost the same as our calibrated economy except that at the beginning housing price equals to the fundamental value. We compare this bubbleless economy with the bubbly economy. In Figure 3.5, we compare the dynamic paths of capital and infrastructure in this two economies. Relative to the bubbleless economy in 2013, capital in bubbly economy is lower in 8.5 percent while infrastructure is higher in 6 percent. The crowding-out effect lowers the output excluding housing in 4 percent while the infrastructure effect highers the output excluding housing in 0.6 percent. Finally, the output excluding housing in bubbly economy is 3.3 percent lower than that in bubbleless economy and the output including

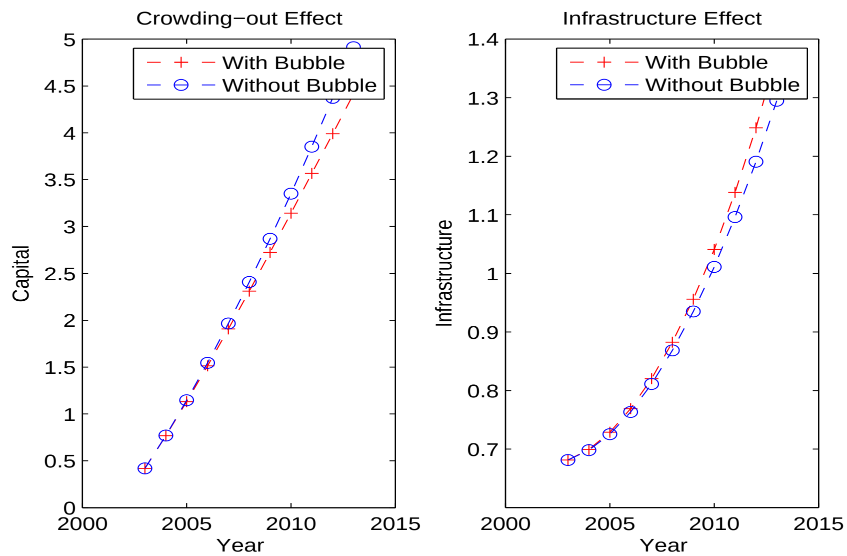

Figure 3.5: Two effects on GDP.

housing is lower in 1 percent.

3.5.2 Bubble Burst

In this subsection, we consider bubble bursts in 2017 and see how bubble burst affects the economy.

First, Figure 3.6 illustrates how bubble burst affects housing price. After bubble bursts, the housing price immediately drops in 40 percent to fundamental value, and the growth rate of housing price changes from 6 percent to -37 percent. Bubble burst does not only change the temporary price but also lowers the long-term growth rate. In the next 30 years after bubble burst, the average growth rate of housing price is 1.9 percent, which is lower than the average growth rate without burst 2.9 percent. This is due to the low rent growth rate. On the balance growth path, growth of housing price with bubble burst is 2.6 percent, which is lower than 4 percent the one without bubble burst.

Second, in Figure 3.7, we show how bubble burst affects GDP. After bubble bursts, the

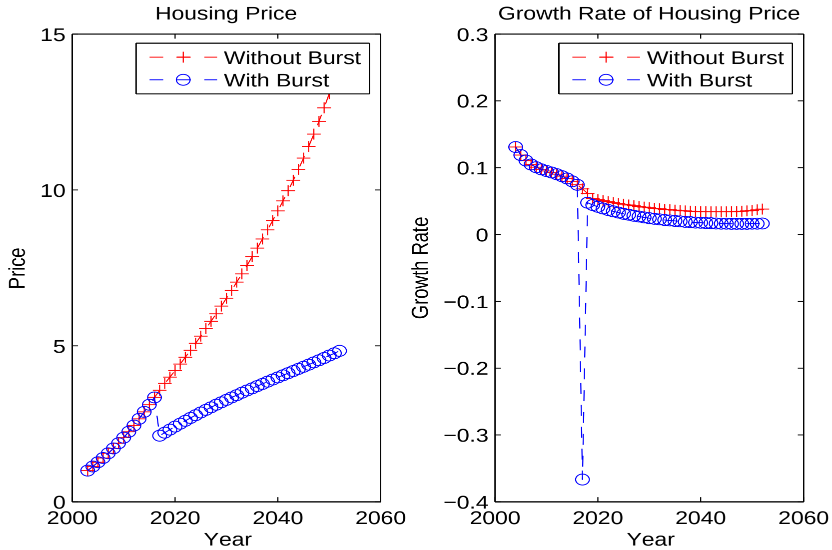

Figure 3.6: Housing price with and without burst.

GDP drops 3.5 percent comparing the level without bubble burst and the GDP growth rate in 2013 drops from 6.2 percent to 2.5 percent. This is due to that the value of newly built house takes in around 10 percent of GDP. When the price drops, newly built house looses value and causes the GDP to drop. However, after the first year of bubble burst, the growth rates of GDP with burst are higher than the ones without burst. The average growth rate with burst in next twenty year is 0.5 percent higher than the average one without burst. Till 2036, twenty years after bubble burst, the GDP with burst is over the GDP without burst for the first time. In 2047, GDP with burst is 1.3 percent higher than GDP with burst and the output excluding housing is 5 percent higher than the one without burst.

The reason for that GDP with burst is higher than the GDP without burst in long run is because infrastructure is over-accumulated in bubbly economy. In 2017, the ratio of infrastructure stock to capital stock is 1:3.8, which is higher than the ratio of production elasticity of infrastructure to the elasticity of capital, which is 1:4.5. The ratio between two elasticities implies the ratio of optimal stocks of infrastructure and capital, when the

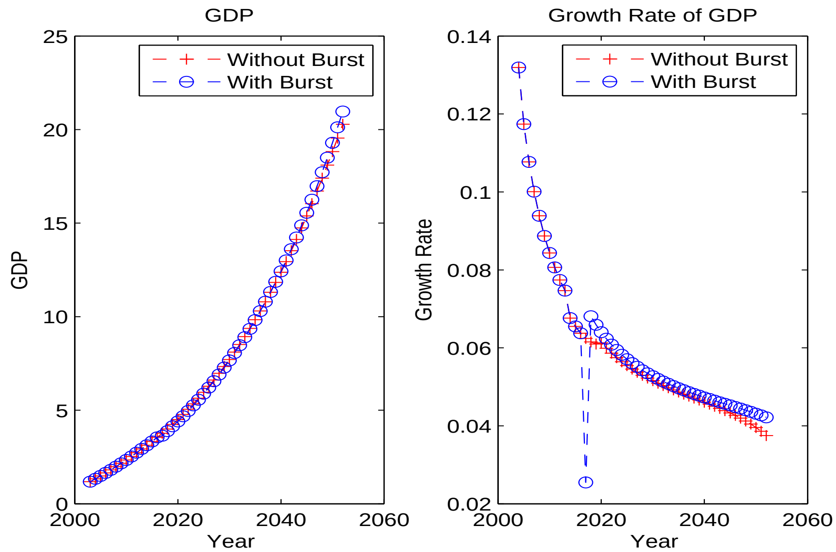

Figure 3.7: GDP with and without burst.

transformation rate from capital to infrastructure is 1:1. In our model, the largest marginal transformation rate is 1:1, which implies even under a conservative consideration, Chinese infrastructure is over-accumulated. Moreover, we find that in 2047, 30 years after bubble bursts, the ratio of infrastructure to capital is 1:6.5, which suggests in a bubbleless economy the infrastructure investment is insufficient. To solve this problem, we suggest a property tax on housing. We will give more details in next subsection.

Bubble burst also causes a huge wealth effect. In Figure 3.8, we show wealth loss for all cohorts in 2017 economy. The cohort born in 1946 lose 16 percent of their wealth due to bubble burst. The younger cohort suffers more due to bubble burst. For the cohort born in 1994, their total wealth lose in 52 percent. The wealth loss for younger cohort comes from not only temporal drop in housing price but also the continuous lower capital return rate. The cohort born in 1995 lose less than the cohort born in 1994, because when bubble bursts they do not hold any housing, thus their wealth loss only comes from the low capital return.

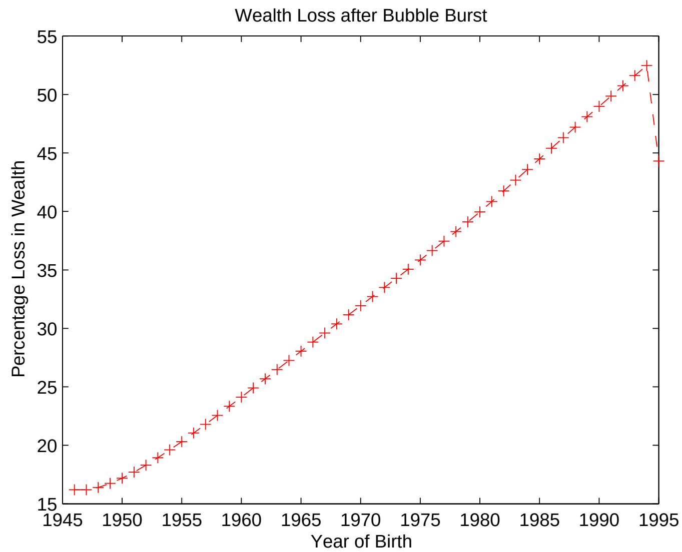

Figure 3.8: Wealth effect of bubble burst.

3.5.3 Property Tax

Chinese government hasn't taken comprehensive property tax so far. In this subsection, we test a case in which Chinese government tax housing with a constant proportion of the total value of housing stock. In our benchmark test, the tax rate is set as 1.5 percent and the tax starts at 2017.

In Figure 3.9, we illustrate how housing price changes after property tax is imposed. The price of housing drops in 59 percent immediately comparing with the price without property tax and the growth rate drops from 6.8 percent to -56 percent. The reason for the drop of price is because former price is not an equilibrium price anymore. If the price does not change, since the growth rate of housing price becomes higher now, total value of housing would eventually be over the size of economy. Because of this, public have to lower their expectation of housing price, and the price drops. One thing interesting is that the drop of housing price when property tax is imposed is larger than the drop when bubble

Figure 3.9: Housing price with and without property tax.

bursts. It is because after imposing property tax, the fundamental value also decreases. After the price immediately drops, in the following ten years, the average growth rate are lower in 0.3 percent than that without property tax, however, in the long term, the growth rate of housing price converges to long-term economy growth rate, same as the one without property tax.

In Figure 3.10, we show how the property tax changes GDP. Based on our simulation. after the property tax is imposed, the GDP drops immediately in 5 percent comparing with the case without property tax.In the growth rate of GDP drops from 6.2 percent to 0.9 percent. However, the long-term output level with property tax is higher than that without property tax. Till 2032, 14 years after property tax started, the GDP level with property tax is higher than that without property tax. In 2047, thirty years after property tax started, the GDP is 2.7 percent higher than that without property tax and the output excluding housing is 6.2 percent higher than that without property tax. One thing needed to be mentioned is that the output level with property tax is higher than the output level with bubble burst.

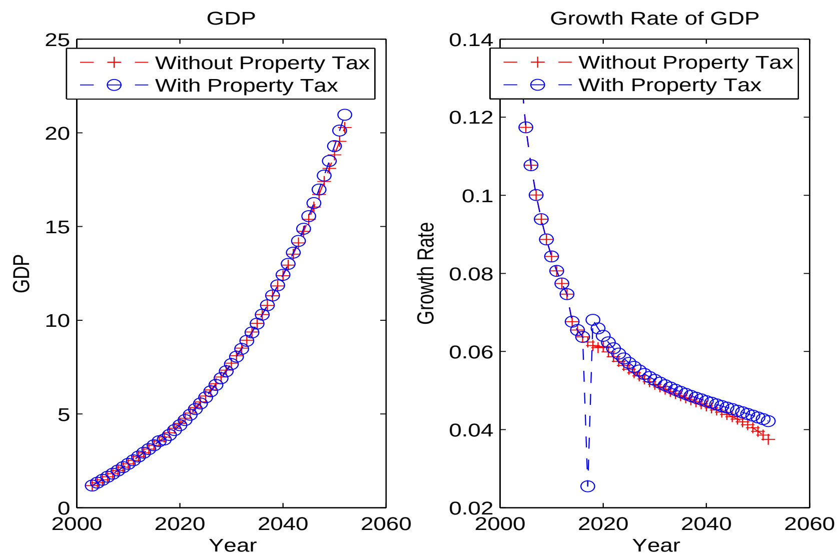

Figure 3.10: GDP with and without property tax.

It is because property tax helps government get fund to accumulate infrastructure. In 2047, infrastructure with property tax is 17 percent higher than that with bubble burst, while capital with property tax is lower only in 0.3 percent than that with bubble burst. Property tax not only lowers the crowding-out effect by lowering the housing price but also compensates for infrastructural investment with new tax.

3.6 Conclusion

In this paper, we study China's housing market in a rational bubble model framework. We view our analysis as making three contributions. First, by allowing a probability of bubble burst, our model can simultaneously account for the steady growth rate of housing price from 2003 to 2013, and the declining rate of return of capital. Second, we quantify the effects of a bubble burst, and find that, although the crash represents a big negative shock to investors' wealth, the effect on China's real GDP is relatively small. The main intuition is that housing market crash would not spread to the rest of the economy. Third, we recommend the adoption of property tax because property tax will 1) reduce housing price even in the absence of a bubble burst, and 2) make up for the lost land-sale revenue that the government needs for infrastructure investment.

Our analysis can be extended in two ways. First, we need to study concave utilities for the entrepreneurs. Under concave utilities, entrepreneurs' investment and consumption decisions are endogenized, and the model can be used to study the impact of lower wealth on entrepreneurs' investment in capital. Opening up this channel may amplify the negative effects of bubble burst on real GDP. Second, we can study the effects of other policy reforms on China's housing market, such as property-purchase limitations and higher down-payment ratios. These extensions are left for future research.

4. SUMMERY

In this dissertation, I applies dynamic method studying two important public issues;one is optimal design of Pacific Salmon Treaty; the other is Chinese housing price bubble.

On the Pacific Salmon Treaty, We extend the theory of continuous-time repeated games to allow for side payments. We view our analysis as making two contributions. First, we show that there are only two solutions to the optimality equation that characterizes the boundary of the set of equilibrium payoffs. This technical contribution greatly simplifies the calculation of the equilibrium set. In the literature, the search for the solution to the optimality equation is done by trial and error. Second, we provide a useful policy recommendation to improve upon the Pacific Salmon Treaty. Because the U.S. production function is more efficient, our optimal contract would give a bigger salmon share to the U.S. than the current treaty does. This policy change will improve the two countries' welfare by 1.54%.

On the Chinese housing price, we study it in a rational bubble model framework. We view our analysis as making three contributions. First, by allowing a probability of bubble burst, our model can simultaneously account for the steady growth rate of housing price from 2003 to 2013, and the declining rate of return of capital. Second, we quantify the effects of a bubble burst, and find that, although the crash represents a big negative shock to investors' wealth, the effect on China's real GDP is relatively small. The main intuition is that housing market crash would not spread to the rest of the economy. Third, we recommend the adoption of property tax because property tax will 1) reduce housing price even in the absence of a bubble burst, and 2) make up for the lost land-sale revenue that the government needs for infrastructure investment.

REFERENCES

[1] J. Shelton and J. P. Koenings, "Marine factors in the production of salmon: Their significance to the pacific salmon treaty," *Alaska Fishery Research Bulletin*, vol. 2, pp. 156–163, 1995.

[2] Y. Sannikov, "Games with imperfectly observable actions in continuous time," *Econometrica*, vol. 75(5), pp. 1285–1329, 2007.

[3] D. Abreu, D. Pearce, and E. Staccheti, "Toward a theory of discounted repeated games with imperfect monitoring," *Econometrica*, vol. 58(5), pp. 1041–1063, 1990.

[4] G. Baker, R. Gibbons, and K. J. Murphy, "Relational contracts and the theory of the firm," *Quarterly Journal of Economics*, vol. 117, pp. 39–84, 2002.

[5] J. Levin, "Relational incentive contracts," *American Economic Review*, vol. 93, pp. 835–857, 2003.

[6] L. Rayo, "Relational incentives and moral hazard in teams," *Review of Economic Studies*, vol. 74, no. 3, pp. 937–963, 2007.

[7] S. Goldlucke and S. Kranz, "Infinitely repeated games with public monitoring and monetary transfers," *Journal of Economic Theory*, vol. 147, no. 3, pp. 1191–1221, 2012.

[8] H. S. Gordon, "The economic theory of a common-property resource: The fishery," *Journal of Political Economy*, vol. 62, pp. 124–142, 1954.

[9] K. Lancaster, "The dynamic inefficiency of capitalism," *Journal of Political Economy*, vol. 81, pp. 1092–1109, 1973.

[10] D. Levhari and L. J. Mirman, "The great fish war: An example using a dynamic cournot-nash solution," *The Bell Journal of Economics*, vol. 11, pp. 322–334, 1980.

[11] J. F. Reinganum and N. L. Stokey, "Oligopoly extraction of a common property natural resource: The importance of the period of commitment in dynamic games," *International Economic Review*, vol. 26, pp. 161–173, 1985.

[12] A. Haurie and M. Pohjola, "Efficient equilibria in a differential game of capitalism," *Journal of Economic Dynamics and Control*, vol. 11, no. 1, pp. 65–78, 1987.

[13] J. Benhabib and R. Radner, "Joint exploitation of a productive asset: A game theoretic approach," *Working Paper, New York University, C.V. Starr Center Appl. Econ.*, 1988.

[14] N. Stokey, *The Economics of Inaction: Stochastic Control Models with Fixed Costs*. New Jersey: Princeton University Press, 2008.

[15] S. Athey and K. Bagwell, "Optimal collusion with private information," *RAND Journal of Economics*, vol. 32, no. 3, pp. 428–465, 2001.

[16] J. Harrington and A. Skrzypacz, "Collusion under monitoring of sales," *RAND Journal of Economics*, vol. 38, no. 2, pp. 314–331, 2007.

[17] L. Samuelson and E. Stacchetti, "Even up: Maintaining relationships," *Working Paper, Yale Univerisity*, 2016.

[18] R. Tinch, I. Dickie, and B. Lanz, "Costs of iuu fishing in eu fisheries," *Technical Report, Economics for the Environment Consultancy*, 2008.

[19] M. Aguiar, M. Amador, and G. Gopinath, "Investment cycles and sovereign debt overhang," *Review of Economic Studies*, vol. 76, pp. 1–31, 2009.

[20] A. Alesina and G. Tabellini, "A positive theory of fiscal deficits and government debt in a democracy," *Review of Economic Studies*, vol. 57, pp. 403–414, 1990.

[21] U. R. Sumaila, “Cooperative and non-cooperative exploitation of the arcto-norwegian cod stock,” *Environmental and Resource Economics*, vol. 10, pp. 147–165, 1997.

[22] S. Villasante and R. Sumaila, “Estimating the economic benefits of cooperative and non-cooperative management of the illex argentinus fishery in south america,” *Working Paper, Latin American and Caribbean Environmental Economics Program*, 2009.

[23] T. Bjorndal and M. Lindroos, “Cooperative and non-cooperative management of the northeast atlantic cod fishery,” *Journal of Bioeconomics*, vol. 14, pp. 41–60, 2012.

[24] K. Tokunaga, “Cooperative management of trans-boundary fish stocks: Implications for tropical tuna management in the pacific island region,” *Working Paper, University of Hawaii*, 2015.

[25] A. Thakur, “Continuous-time principal multi-agent problem: Moral hazard in teams & fiscal federalism,” *Working Paper, Princeton Univerisity*, 2015.

[26] J. Tirole, “Asset bubbles and overlapping generations,” *Econometrica*, vol. 53, no. 6, pp. 1499–1528, 1985.

[27] K. Chen and Y. Wen, “The great housing boom of china,” *Working Paper, Federal Reserve Bank of St. Louis*, 2015.

[28] H. Fang, Q. Gu, W. Xiong, and L.-A. Zhou, “Demystifying the chinese housing boom,” in *NBER Macroeconomics Annual 2015, Volume 30*, pp. 105–166, National Bureau of Economic Research, Inc, 2015.

[29] N. Kiyotaki and J. Moore, “Credit cycles,” *Journal of Political Economy*, vol. 105, no. 2, pp. 211–48, 1997.

[30] J. Miao and P. Wang, “Bubbles and credit constraints,” Boston University - Department of Economics - Working Papers Series WP2011-031, Boston University - Department of Economics, 2011.

[31] A. Martin and J. Ventura, “Economic growth with bubbles,” *American Economic Review*, vol. 102, no. 6, pp. 3033–58, 2012.

[32] E. Farhi and J. Tirole, “Bubbly liquidity,” *Review of Economic Studies*, vol. 79, no. 2, pp. 678–706, 2012.

[33] Z. Song, K. Storesletten, and F. Zilibotti, “Growing like china,” *American Economic Review*, vol. 101, no. 1, pp. 196–233, 2011.

[34] R. Barro, “Output effects of government purchases,” *Journal of Political Economy*, vol. 89, no. 6, pp. 1086–1121, 1981.

[35] R. Barro, “Government spending in a simple model of endogenous growth,” *Journal of Political Economy*, vol. 98, no. 5, pp. S103–26, 1990.

[36] M. Baxter and R. King, “Fiscal policy in general equilibrium,” *American Economic Review*, vol. 83, no. 3, pp. 315–34, 1993.

[37] G. Glomm and B. Ravikumar, “Public investment in infrastructure in a simple growth model,” *Journal of Economic Dynamics and Control*, vol. 18, pp. 1173–1187, 1994.

[38] M. Bassetto and T. J. Sargent, “Politics and efficiency of separating capital and ordinary government budgets,” *The Quarterly Journal of Economics*, vol. 121, no. 4, pp. 1167–1210, 2006.

[39] C.-E. Bai, C.-T. Hsieh, and Y. Qian, “The return to capital in china,” *Brookings Papers on Economic Activity*, vol. 37, no. 2, pp. 61–102, 2006.

[40] P. Bom and J. Ligthart, "What have we learned from three decades of research on the productivity of public capital?," *Journal of Economic Surveys*, vol. 28, no. 5, pp. 889–916, 2014.

[41] G. Jin, "Infrastructure and non-infrastructure capital stocks in china and their productivityïijŽa new estimate," *Economic Research Journal*, no. 5, pp. 41–56, 2016.

[42] C.-E. Bai and Z. Qiong, "Return to capital in china and its determinants," *Brookings Papers on Economic Activity*, vol. 48, no. 10, pp. 3–30, 2014.

[43] N. Portley, C. Hendrich, and K. Balliet, "Sfp global sustainability overview of pacific salmon fisheries," *Technical Report, Sustainable Fisheries Partnership Foundation*, 2013.

[44] N. Portley, P. Sousa, B. Lee-Harwood, C. Hendrich, and K. Balliet, "Global sustainability overview of pacific salmon fisheries," *Technical Report, Sustainable Fisheries Partnership Foundation*, 2014.

[45] C. Stroomer and M. Wilson, *British Columbia's Fisheries and Aquaculture Sector*. Victoria, BC, Canada: BC STATS, 2012 ed., 2013.

[46] S. Carlson, "2002 survey of bristol bay salmon drift gillnet fishery permit holders: Preliminary summary of response," *Technical Report, CFEC*, 2002.

[47] K. Schelle, K. Iverson, N. Free-Sloan, and S. Carlson, "Bristol bay salmon drift gillnet fishery optimum number report," *Technical Report, CFEC*, 2004.

[48] D. S. Liao, "Profitability and productivity analysis for the southeastern alaska salmon fishery," *Marine Fisheries Review*, vol. 38, pp. 11–14, 1976.

[49] D. D. Huppert, G. M. Ellis, and B. Noble, "Do permit prices reflect the discounted value of fishing? evidence from alaska's commercial salmon fisheries," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 53, no. 4, pp. 761–768, 1996.

[50] P. Hartman, *Ordinary Differential Equations*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2nd ed., 2002.

Lemma Corollary Definition Assumption Remark Proposition Theorem

APPENDIX A

APPENDIX OF SECTION TWO

Appendix A.1: A brief history of the Pacific Salmon Treaty

The cooperative management of Pacific salmon between the U.S. and Canada can be traced back to the 1930s. In 1930, the two governments signed the Fraser River Convention to equally share fish resources in the transboundary Fraser River. During the 1970s, interception became a major problem undermining the two countries' cooperation. Fishing vessels from British Columbia caught many coho and chinook originating in Washington, while vessels from Alaska harvested a lot of sockeye originating in British Columbia. Solving this problem required that the two countries seek an agreement along the entire west coast. In 1985, after 14 years of negotiation, they signed the initial version of the Pacific Salmon Treaty. Based on this treaty, the two countries formed the Pacific Salmon Commission to design fishing plans (fishing amount, fishing schedule, etc.) for both countries. However, after the initial version expired in 1992, they could not reach any new agreement due to interception disputes. Consequently, there was no cooperative management of Pacific salmon from 1992 to 1998. This period, during which numerous conflicts occurred, is commonly referred to as one of fish war.

The U.S. and Canada signed a new version of the Pacific Salmon Treaty in 1999. The new treaty has two features. First, its sharing rule is dynamically adjusted. Take sockeye in the Fraser River for example.[1] In 1999, the pre-season fishing plan gave the U.S. 22.4% of the total allowable catch of sockeye. In 2000, this number was adjusted to 20.4%. The adjustment was based on the sharing rule in the treaty, which took into account historic

[1] Sharing rules in several other transboundary rivers (such as the Stikine, Taku, and Alsek rivers) are dynamically adjusted too.

catch and other ecological data. Second, the new treaty allows one country to compensate the other by making side-payments. In 1999, the U.S. made an implicit transfer to Canada by providing $140 million to establish two restoration and enhancement funds ("Northern Fund" and "Southern Fund"). The majority of the funds are spent on projects that enhance Canadian salmon even if Canada did not supply the funds. In 2009, the U.S. transferred $30 million to Canada to compensate for Canada's loss from its fishery mitigation project. The goal of this project is to downsize some Canadian fisheries.

The new Pacific Salmon Treaty has been well enforced. Fishing disputes similar to those during the 1992-1998 fish war have not occurred since 1999.[2] Moreover, data from the U.S. National Oceanic and Atmospheric Administration (NOAA) show that overfishing is no longer a serious issue in the salmon industry. More specifically, we use NOAA's data to compute the fraction of the overfished salmon stock within the total salmon stock, and find that this fraction is small after 2000 (see Figure A.1).[3] This evidence suggests that the overall stock of salmon has been stabilized.[4] Hence, in this paper, we assume for simplicity that the total stock of salmon is constant over time.

Appendix A.2

A.2.1 The ratio of value added to revenue

For Canada, [45] provide data on the annual revenue and value added of Pacific salmon from 1990 to 2011. We calculate the ratio of value added to revenue in each year, and then average the ratios to get 22.5%. Since [45] is a report prepared for Canada's Department

[2]The Pacific Salmon Commission, which carries out the Pacific Salmon Treaty, has never had its normal activity disrupted since 1999. The Commission holds three bilateral meetings annually, and its next scheduled meeting is September 25-29, 2017 in Portland, Oregon. See http://www.psc.org/meetings/schedule/.

[3]NOAA assesses nationwide fish stocks of different species and areas every quarter, and uses information such as fisheries landings, scientific surveys, and biological studies to determine whether a stock is overfished.

[4]Unfortunately, we could not find complete time-series data about Canada's Pacific salmon stock. Nevertheless, [43] and [44] find that 96% and 95% of Pacific salmon in British Columbia were well managed in 2013 and 2014, respectively.

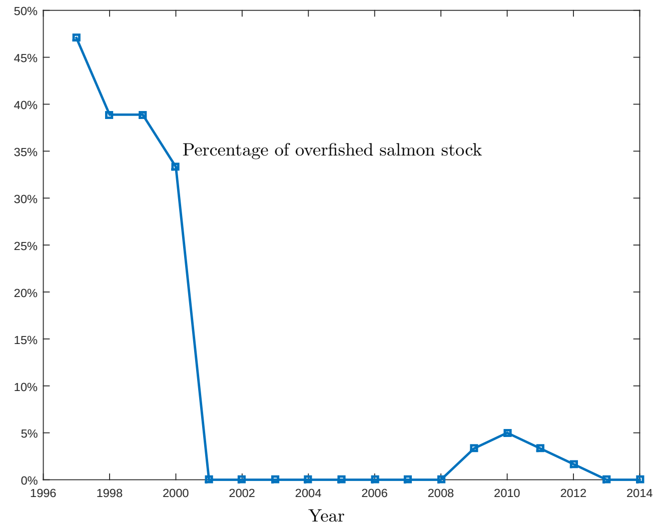

Figure A.1: Fraction of the U.S. Pacific salmon stock in overfished status.

of Fisheries and Oceans, this data source is reliable.

For the U.S., we have not found nationwide data on value added of Pacific salmon. Below is a list of various estimates based on regional data.

1. [46] contains the revenue and different categories of costs based on a 2002 survey on Pacific salmon fishing in Bristol Bay. From the data, we calculate the value-added ratio as 40.9%. Because salmon production in Bristol Bay is around one seventh of the total U.S. production, we think these data are representative.

2. [47] construct estimators of the cost and revenue of salmon catching in Bristol Bay, using multiple data sources including the 2002 survey in (i). Based on their estimators, the average value-added ratio was 45.8% from 1999 to 2003.

3. [48] provides the revenue and costs from a 1973 survey in the Southeastern Alaska Salmon Fishery. From these data, we calculate the value-added ratio as 48.1%.

4. [49] provide data from a 1979 survey of all salmon fisheries in Alaska. The (incomplete) data suggest that the value-added ratio is at least 46%.

All the above estimates of the U.S. value-added ratio are higher than Canada's 22.5%. This suggests that allocating more salmon to the U.S. will improve total welfare. In order to be conservative in estimating the welfare gain, we choose the smallest from the above list (i.e., 40.9% in (i)) as the ratio for the U.S. in Subsection 3.2.1.

It is not surprising that the U.S. has a higher value-added ratio in salmon than Canada. For the whole fishery industry, the average value-added ratios during 1999-2012 for the U.S. and Canada were, respectively, 63.3% and 49.7%, and the ratio for British Columbia was only 37.0%.[5]

A.2.2 Robustness check

In this subsection, we check the robustness of our results with respect to certain parameters. First, we try different values for the parameter of the cost of illegal fishing, μ. In our benchmark calibration, we set $\mu = 0.84$ using data from the European Union, but the cost of illegal fishing in North America may be either higher or lower than 0.84. It turns out that our numerical results are insensitive to μ. In particular, the welfare gains are 1.59% and 1.54%, respectively, under a lower $\mu = 0.5 * 0.84$ and a higher $\mu = 2 * 0.84$, as opposed to 1.54% under the benchmark $\mu = 0.84$.

Second, we redo the calibration and recalculate the welfare gain using a higher estimate of the U.S. value-added ratio. We choose the highest estimate of 48.1% from (iii) in Section B.1. Our results are reported in Table A.1.

The welfare gain of $92.51 million (or 2.37%) is larger than the $51.66 million (or 1.54%) in Table 2.4. This larger welfare gain is mainly because a more efficient U.S. production function has made it more profitable to reallocate resources. To see this more

[5]Data sources: Fisheries of the United States (1999-2012) and Statistics Canada.

Table A.1: Change of welfare measured in 1999 million USD when the U.S. value-added ratio is 48.1%

	Positive contract	Optimal contract	Difference
$E\left[\int_0^\infty e^{-rt}(p_{1t}+p_{2t})\,dt - \sum_{k=1}^\infty e^{-rt_k} C\right]$	3902.62	3995.13	92.51
$E\left[\sum_{k=1}^\infty e^{-rt_k} C\right]$	46.43	12.35	-34.08
$E\left[\int_0^\infty e^{-rt}(p_{1t}+p_{2t})\,dt\right]$	3949.04	4007.48	58.44
$E\left[\int_0^\infty e^{-rt} p_{1t} dt\right]$	3794.5	4000.19	205.69
$E\left[\int_0^\infty e^{-rt} p_{2t} dt\right]$	154.55	7.3	-147.25

clearly, note that the welfare gain of \$58.44 million due to the increase in $p_{1t} + p_{2t}$ is much larger than the \$34.08 million due to the reduction of fixed cost, whereas in Table 2.4 the two welfare gains are close. The increase in the present value of $p_{1t} + p_{2t}$ in Table A.1 is again due to reallocating resources from Canada to the U.S., only on a bigger scale here. The average catching weight of Canada has decreased to 0.81 million pounds in the optimal contract in Table A.1, as opposed to 2.91 million pounds in Table 2.4. This explains why Canada's present value of p_{2t} is \$7.3 million, much lower than the \$22.91 million in Table 2.4.

Appendix A.3: Proofs

Proof of Lemma 2: If an IC contract $(\mathbf{x}, \mathbf{e}, \mathbf{Q})$ with continuation payoff process $\mathbf{W}$ satisfies $e_{it} > 0$, then we can define an equivalent contract $(\tilde{\mathbf{x}}, \tilde{\mathbf{e}}, \tilde{\mathbf{Q}})$ with identical payoff $\tilde{\mathbf{W}} = \mathbf{W}$, but without stealing. In particular, define

$$\tilde{x}_t := x_t,$$

$$\tilde{e}_{it} := 0, \qquad \tilde{e}_{jt} := e_{jt},$$

$$\tilde{Q}_{it} := Q_{it}, \quad d\tilde{Q}_{jt} := dQ_{jt} + e_{it}dt.$$

In the new contract, $\tilde{W}_{it} = W_{it}$ because player i is indifferent between stealing $e_{it}dt$ from his opponent and being paid $e_{it}dt$ by his opponent. To guarantee $\tilde{W}_{jt} = W_{jt}$, the new contract may burn $(\mu - \tau)e_{it}dt$ units of the resource to increase player j's total payment expense at time t to $(1+\mu)e_{it}dt = (1+\tau)e_{it}dt + (\mu-\tau)e_{it}dt$. Then player j is indifferent between incurring payment expense $(1+\mu)e_{it}dt$ and having $(1+\mu)e_{it}dt$ being stolen by player i.

To finish the proof, we verify that $(\tilde{\mathbf{x}}, \tilde{\mathbf{e}}, \tilde{\mathbf{Q}})$ remains IC. It follows from $\tilde{\mathbf{W}} = \mathbf{W}$ that $(\tilde{\beta}^{i1}, \tilde{\beta}^{i2}) = (\beta^{i1}, \beta^{i2})$, $i = 1, 2$. Therefore the IC constraint $1 + \tilde{\beta}^{ii} - (1+\mu)\tilde{\beta}^{ij} \leq 0$ follows from (2.3).

Proof of Lemma 3:

1. A relaxed problem in which player 2's incentive constraint is removed is

$$\min_{\phi_1,\phi_2} \quad \sigma_1^2\phi_1^2 + \sigma_2^2\phi_2^2$$
$$\text{subject to} \quad 1 \leq \sin(\theta)(\phi_1 + (1+\mu)\phi_2).$$

The optimal solution is

$$\phi_1 = \frac{\sigma_2^2}{\sin(\theta)(\sigma_2^2 + (1+\mu)^2\sigma_1^2)}, \quad \phi_2 = \frac{(1+\mu)\sigma_1^2}{\sin(\theta)(\sigma_2^2 + (1+\mu)^2\sigma_1^2)},$$

which implies

$$\sigma_1^2\phi_1^2 + \sigma_2^2\phi_2^2 = \frac{\sigma_1^2\sigma_2^2}{(\sigma_2^2 + (1+\mu)^2\sigma_1^2)\sin^2(\theta)}.$$

Player 2's incentive constraint is slack (and player 1's incentive constraint is binding) when

$$1 \leq \cos(\theta)((1+\mu)\phi_1 + \phi_2) = \frac{\cos(\theta)}{\sin(\theta)}\frac{(\sigma_1^2 + \sigma_2^2)(1+\mu)}{(\sigma_2^2 + (1+\mu)^2\sigma_1^2)},$$

or

$$\frac{\sin(\theta)}{\cos(\theta)} \leq \frac{(\sigma_1^2 + \sigma_2^2)(1+\mu)}{\sigma_2^2 + (1+\mu)^2\sigma_1^2}.$$

Similarly, player 1's incentive constraint is slack (and player 2's incentive constraint is binding) if

$$\frac{\sin(\theta)}{\cos(\theta)} \geq \frac{\sigma_1^2 + (1+\mu)^2\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)(1+\mu)}.$$

If $\frac{\sin(\theta)}{\cos(\theta)} \in \left[\frac{(\sigma_1^2+\sigma_2^2)(1+\mu)}{\sigma_2^2+(1+\mu)^2\sigma_1^2}, \frac{\sigma_1^2+(1+\mu)^2\sigma_2^2}{(\sigma_1^2+\sigma_2^2)(1+\mu)}\right]$, both constraints bind. The optimal solution is

$$\phi_1 = \frac{\frac{1+\mu}{\cos(\theta)} - \frac{1}{\sin(\theta)}}{(1+\mu)^2 - 1}, \quad \phi_2 = \frac{\frac{1+\mu}{\sin(\theta)} - \frac{1}{\cos(\theta)}}{(1+\mu)^2 - 1}.$$

This implies

$$\sigma_1^2\phi_1^2 + \sigma_2^2\phi_2^2 = \frac{a - b\sin(2\theta) + c\cos(2\theta)}{\sin^2(2\theta)},$$

where

$$
\begin{aligned}
a &:= \frac{2((1+\mu)^2+1)(\sigma_1^2+\sigma_2^2)}{((1+\mu)^2-1)^2}, && \text{(A.1)}\\
b &:= \frac{4(1+\mu)(\sigma_1^2+\sigma_2^2)}{((1+\mu)^2-1)^2}, && \text{(A.2)}\\
c &:= \frac{2(\sigma_2^2-\sigma_1^2)}{(1+\mu)^2-1}. && \text{(A.3)}
\end{aligned}
$$

2. The optimality equation can be rewritten as

$$
\begin{aligned}
\kappa(w) &= \max_{x\in[0,1]} \frac{2\mathbf{N}(\theta)((p_1(x),p_2(x))-w)}{r|\phi(\theta)|^2}\\
&= \frac{\max_{x\in[0,1]} 2(\cos(\theta)p_1(x)+\sin(\theta)p_2(x)) - 2\mathbf{N}(\theta)w}{r|\phi(\theta)|^2}.
\end{aligned}
$$

So the optimal x^* satisfies the first-order condition

$$\cos(\theta)p_1'(x^*)+\sin(\theta)p_2'(x^*)=0. \tag{A.4}$$

Since both $p_1(x)$ and $p_2(x)$ are concave functions, the above equation has a unique solution x^*.

Proof of Lemma 4: First, the definition of $\overline{x}$ implies that $(1+\tau)p_1(\overline{x})+p_2(\overline{x}) \geq (1+\tau)p_1(x)+p_2(x)$ for all $x \in [0,1]$. Second, the definition of continuation payoffs in (2.1) implies

$$
\begin{aligned}
& (1+\tau)W_1 + W_2 \\
= \; & E\left[\int_0^\infty re^{-rt}\left(((1+\tau)p_1(x_t)+p_2(x_t))dt - r(1+\tau)^2 dQ_{1t} + rdQ_{1t}\right)\right] \\
\leq \; & E\left[\int_0^\infty re^{-rt}\left((1+\tau)p_1(\overline{x})+p_2(\overline{x})\right)dt\right] \\
= \; & (1+\tau)p_1(\overline{x})+p_2(\overline{x}).
\end{aligned}
$$

If $(W_1, W_2 = \underline{u}_2)$ is the promised payoff of some contract, then $(1+\tau)W_1 + \underline{u}_2 \leq (1+\tau)p_1(\overline{x})+p_2(\overline{x})$, which implies $W_1 \leq \overline{\overline{u}}_1$.

Lemma A.1. *The right-hand sides of* (2.10)-(2.12) *satisfy the Lipschitz condition in the open set $\mathcal{B}$ defined by*

$$
\begin{aligned}
\mathcal{B} := \Big\{(w_1, w_2, \theta, l) : \quad & \underline{u}_1 - \epsilon < w_1 < \overline{\overline{u}}_1 + \epsilon, \\
& \underline{u}_2 - \epsilon < w_2 < \underline{u}_2 + \frac{\overline{\overline{u}}_1 - \underline{u}_1 + 2\epsilon}{\tan(\underline{\theta})}, \\
& 0 < \theta < \frac{\pi}{2}, \quad -\infty < l < \infty\Big\},
\end{aligned}
$$

where $\epsilon > 0$ is a small positive number.

Proof: Equations (2.10) and (2.11) satisfy the Lipschitz condition because their derivatives with respect to θ are bounded. To show the Lipschitz continuity of (2.12), define

$$
\begin{aligned}
m(\theta) \; &:= \; \frac{1}{|\phi(\theta)|^2}, \\
n(\theta, w) \; &:= \; \frac{2}{r}\max_{x\in[0,1]} \mathbf{N}(\theta)((p_1(x), p_2(x)) - w) = \frac{2}{r}\mathbf{N}(\theta)((p_1(x^*), p_2(x^*)) - w),
\end{aligned}
$$

where x^* is the optimal strategy given in (A.4). So the right-hand side of (2.12) is equal to $m(\theta)n(\theta, w)$. The rest of the proof consists of three steps.

First, both m and n are bounded and continuous. Function n is bounded because $\mathbf{N}(\theta)$, $p_1(x^*)$, $p_2(x^*)$, and w are all bounded in $\mathcal{B}$; it is continuous because x^* is continuous in θ. To see that m is bounded and continuous for $\theta \in (0, \frac{\pi}{2})$, recall from Lemma 3 that

$$m(\theta) = \begin{cases} (\sigma_2^{-2} + (1+\mu)^2\sigma_1^{-2})\sin^2(\theta), & \text{if } \theta \in (0, \theta_1]; \\ \frac{\sin^2(2\theta)}{a-b\sin(2\theta)+c\cos(2\theta)}, & \text{if } \theta \in [\theta_1, \theta_2]; \\ (\sigma_1^{-2} + (1+\mu)^2\sigma_2^{-2})\cos^2(\theta), & \text{if } \theta \in [\theta_2, \frac{\pi}{2}), \end{cases} \tag{A.5}$$

where a, b, and c are defined in (A.1)-(A.3). Function m is bounded on $(0, \theta_1] \cup [\theta_2, \frac{\pi}{2})$ because $m \leq \max(\sigma_2^{-2} + (1+\mu)^2\sigma_1^{-2}, \sigma_1^{-2} + (1+\mu)^2\sigma_2^{-2})$; it is bounded on $[\theta_1, \theta_2]$ because $\frac{\sin^2(2\theta)}{a-b\sin(2\theta)+c\cos(2\theta)} \leq \frac{1}{a-\sqrt{b^2+c^2}}$.[6] Function m is continuous because Lemma 2 shows that both ϕ_1 and ϕ_2 are continuous in θ.

Second, $m(\theta)n(\theta, w)$ is Lipschitz continuous in w because the partial derivative of $m(\theta)n(\theta, w)$ with respective to w is bounded. In particular,

$$\frac{\partial(m(\theta)n(\theta,w))}{\partial w_1} = -\frac{2}{r}m(\theta)\cos(\theta), \quad \frac{\partial(m(\theta)n(\theta,w))}{\partial w_2} = -\frac{2}{r}m(\theta)\sin(\theta),$$

which are bounded because m is bounded.

Third, $m(\theta)n(\theta, w)$ is Lipschitz continuous in θ. To show this, it is sufficient to show that $\frac{\partial(m(\theta)n(\theta,w))}{\partial\theta}$ is bounded. Because $(mn)' = m'n + mn'$ and both m and n are bounded, it is sufficient to show that m' is bounded and n' is bounded.

1. $m'(\theta)$ is bounded. Differentiating (A.5) yields

$$m'(\theta) = \begin{cases} (\sigma_2^{-2} + (1+\mu)^2\sigma_1^{-2})\sin(2\theta), & \text{if } \theta \in (0, \theta_1]; \\ \frac{2\sin(4\theta)(a-b\sin(2\theta)+c\cos(2\theta))+\sin^2(2\theta)(2b\cos(2\theta)+2c\sin(2\theta))}{(a-b\sin(2\theta)+c\cos(2\theta))^2}, & \text{if } \theta \in [\theta_1, \theta_2]; \\ -(\sigma_2^{-2} + (1+\mu)^2\sigma_1^{-2})\sin(2\theta), & \text{if } \theta \in [\theta_2, \frac{\pi}{2}). \end{cases}$$

[6] $a - \sqrt{b^2+c^2} > 0$ because $a^2 - (b^2+c^2) = 16\sigma_1^2\sigma_2^2/((1+\mu)^2-1)^2 > 0$.

Function m' is bounded on $(0,\theta_1]\cup[\theta_2,\frac{\pi}{2})$ because $m' \le \max(\sigma_2^{-2}+(1+\mu)^2\sigma_1^{-2}, \sigma_1^{-2}+(1+\mu)^2\sigma_2^{-2})$; it is bounded on $[\theta_1,\theta_2]$ because

$$\begin{aligned} & \left|\frac{2\sin(4\theta)(a-b\sin(2\theta)+c\cos(2\theta))+\sin^2(2\theta)(2b\cos(2\theta)+2c\sin(2\theta))}{(a-b\sin(2\theta)+c\cos(2\theta))^2}\right| \\ \le \quad & \frac{2(|a|+|b|+|c|)+(2|b|+2|c|)}{(a-\sqrt{b^2+c^2})^2}. \end{aligned}$$

Hence $m'(\theta)$ is bounded.

2. $n'(\theta)$ is bounded. The envelope theorem implies

$$n'(\theta) = \frac{2}{r}\left(-\sin(\theta)(p_1(x^*)-w_1)+\cos(\theta)(p_2(x^*)-w_2)\right),$$

where x^* is the optimal strategy in (A.4). Because w is bounded in $\mathcal{B}$ and $x^* \in [0,1]$, $n'(\theta)$ is also bounded.

Lemma A.2. *1. If a solution to* (2.10)-(2.12) *starts from* $(w_1,w_2,\theta)=(\overline{\overline{u}}_1,\underline{u}_2,\underline{\theta})$*, then it is a straight line.*

2. If a solution to (2.10)-(2.12) *starts from* $(w_1,w_2,\theta)=(\overline{u}_1,\underline{u}_2,\underline{\theta})$*, where* $\overline{u}_1<\overline{\overline{u}}_1$*, then its curvature is positive (i.e.,* $\theta'(l)>0$*) and* $\theta(l)<\frac{\pi}{2}$ *for all* $l\ge 0$.

Proof:

1. Define a straight line by $(w_1(l),w_2(l),\theta(l)) := (\overline{\overline{u}}_1-\sin(\underline{\theta})l, \underline{u}_2+\cos(\underline{\theta})l, \underline{\theta})$. This straight line solves (2.10)-(2.12) because

$$\begin{aligned} \max_{x\in[0,1]} \frac{2\mathbf{N}(\underline{\theta})((p_1(x),p_2(x))-w)}{r|\phi(\underline{\theta})|^2} &= \frac{2\mathbf{N}(\underline{\theta})((p_1(\overline{x}),p_2(\overline{x}))-w)}{r|\phi(\underline{\theta})|^2} \\ &= 0 = \theta'(l). \end{aligned}$$

A solution to (2.10)-(2.12) must equal the above straight line because the solution is unique.

2. First, we prove $\theta'(l) > 0, \forall l \geq 0$ by contradiction. If $\overline{u}_1 < \overline{\overline{u}}_1$, then the optimal equation implies that $\theta'(l = 0) > 0$. Suppose $\theta'(l^*) = 0$ at some $l^* > 0$. Then the straight line that passes through $w(l^*)$ and is parallel to $T(\theta(l^*))$ solves (2.10)-(2.12) from the initial conditions $(w(l^*), \theta(l^*))$.[7] It has zero curvature throughout. However, the original curve $(w(l), \theta(l))$ also solves (2.10)-(2.12), and passes through $w(l^*)$, but has positive curvature at $\theta = \underline{\theta}$. This contradicts the property that the solution to (2.10)-(2.12) is unique.

   Second, we show $\theta(l) < \frac{\pi}{2}, \forall l \geq 0$ by contradiction. Suppose $\theta(l^*) = \frac{\pi}{2}$ at some $l^* > 0$. For any integer k, define l_k by $\theta(l_k) = \frac{\pi}{2} - \frac{1}{2^k}$. Since θ increases in l, $\{l_k\}$ is an increasing sequence. Therefore,

$$\begin{aligned}\frac{1}{2^{k+1}} = \theta(l_{k+1}) - \theta(l_k) &= \int_{l_k}^{l_{k+1}} m(\theta(l))n(\theta(l))dl \\ &= m(\theta(\bar{l}))n(\theta(\bar{l}))(l_{k+1} - l_k) \\ &\leq m(\theta(\bar{l}))N(l_{k+1} - l_k),\end{aligned}$$

   where $\bar{l}$ is a point in $[l_k, l_{k+1}]$ and N is an upper bound for the bounded function n. Because m is shown to be Lipschitz continuous in Lemma A.1, $m(\theta(\bar{l})) \leq M(\frac{\pi}{2} -$

[7] Denote $(w_1(l^*), w_2(l^*), \theta(l^*))$ as (w_1^*, w_2^*, θ^*). As in step (i), the straight line defined by $(\tilde{w}_1(l), \tilde{w}_2(l), \tilde{\theta}(l)) := (w_1^* - \sin(\theta^*)(l - l^*), w_2^* + \cos(\theta^*)(l - l^*), \theta^*)$ satisfies (2.12) because

$$\tilde{\theta}'(l) = 0 = \max_{x \in [0,1]} \frac{2\mathbf{N}(\theta^*)((p_1(x), p_2(x)) - \tilde{w})}{r|\phi(\theta^*)|^2}.$$

$\theta(\bar{l}))$ for some M. Therefore,

$$\begin{aligned}\frac{1}{2^{k+1}} \leq m(\theta(\bar{l}))N(l_{k+1}-l_k) &\leq M(\frac{\pi}{2}-\theta(\bar{l}))N(l_{k+1}-l_k) \\ &\leq MN(\frac{\pi}{2}-\theta(l_k))(l_{k+1}-l_k) \\ &= MN\frac{1}{2^k}(l_{k+1}-l_k),\end{aligned}$$

which implies $l_{k+1} - l_k > \frac{1}{2MN}$ for all k. This contradicts the assumption that $l^* = \lim_{k\to\infty} l_k$ is finite.

Proof of Lemma 5: Because the ODE system in (2.10)-(2.12) satisfies the Lipschitz condition in $\mathcal{B}$, the extension theorem (e.g., [Theorem 3.1, page 12]Hartman2002) states that, starting from any initial condition in $\mathcal{B}$, a unique solution exists and extends to the boundary of $\mathcal{B}$. In particular, if $\overline{u}_1 \in [\underline{u}_1, \overline{\overline{u}}_1]$ and the initial condition is $(w_1, w_2, \theta) = (\overline{u}_1, \underline{u}_2, \underline{\theta})$ and $l = 0$, then the solution extends to $(w_1(l^+), w_2(l^+), \theta(l^+), l^+) \in \partial\mathcal{B}$.

We show that $w_1(l^+) = \underline{u}_1 - \epsilon$. The proof consists of several steps.

1. $l^+ < \infty$. In fact, $l^+ \leq \frac{\overline{u}_1 - \underline{u}_1 + \epsilon}{\sin(\underline{\theta})}$ follows from

$$\begin{aligned}\underline{u}_1 - \epsilon \leq w_1(l^+) = \overline{u}_1 - \int_0^{l^+} \sin(\theta(l))dl &\leq \overline{u}_1 - \int_0^{l^+} \sin(\underline{\theta})dl \\ &= \overline{u}_1 - \sin(\underline{\theta})l^+.\end{aligned}$$

2. $\theta(l^+) < \frac{\pi}{2}$. Because l^+ is finite, part (ii) of Lemma A.2 implies $\theta(l^+) < \frac{\pi}{2}$.

3. $w_1(l^+) < \overline{\overline{u}}_1 + \epsilon$. This follows from $w_1'(l) = -\sin(\theta) < 0$.

4. $\underline{u}_2 - \epsilon < w_2(l^+)$. This follows from $w_2'(l) = \cos(\theta) > 0$.

5. $w_2(l^+) < \underline{u}_2 + \cos(\underline{\theta})\frac{\overline{\overline{u}}_1 - \underline{u}_1 + 2\epsilon}{\sin(\underline{\theta})}$. This follows from

$$\begin{aligned} w_2(l^+) &= \underline{u}_2 + \int_0^{l^+} \cos(\theta(l))dl \leq \underline{u}_2 + \int_0^{l^+} \cos(\underline{\theta})dl \\ &\leq \underline{u}_2 + \cos(\underline{\theta})\frac{\overline{u}_1 - \underline{u}_1 + \epsilon}{\sin(\underline{\theta})} \\ &< \underline{u}_2 + \cos(\underline{\theta})\frac{\overline{\overline{u}}_1 - \underline{u}_1 + 2\epsilon}{\sin(\underline{\theta})}, \end{aligned}$$

where the second inequality follows from step (i), and the third inequality follows from $\overline{u}_1 \leq \overline{\overline{u}}_1$.

6. $w_1(l^+) = \underline{u}_1 - \epsilon$. Otherwise, $(w_1(l^+), w_2(l^+), \theta(l^+), l^+) \notin \partial\mathcal{B}$.

Because $w_1(l^+) = \underline{u}_1 - \epsilon$ and $w_1(0) = \overline{u}_1 \geq \underline{u}_1$, the intermediate value theorem implies the existence of an L such that $w_1(L) = \underline{u}_1$. It follows from $w_1'(l) < 0$ that L is unique.

The following lemmas A.3-A.5 are used in the proof of Theorem 1.

Lemma A.3. Θ *is a continuous function of* $\overline{u}_1 \in [\underline{u}_1, \overline{\overline{u}}_1]$.

Proof: Since this proof considers solution curves of various initial conditions $\overline{u}_1$, we shall write w_1 as a function of both l and $\overline{u}_1$. Because the optimality equation satisfies the Lipschitz condition, its solution $w_1(l, \overline{u}_1)$ is continuous in $(l, \overline{u}_1)$ (e.g., [Theorem 2.1, page 94] [50]).

First, we show the L defined in Lemma 5 is a continuous function of $\overline{u}_1 \in [\underline{u}_1, \overline{\overline{u}}_1]$. Because the point $(L(\overline{u}_1), \overline{u}_1)$ satisfies the equation

$$w_1(l, \overline{u}_1) = \underline{u}_1,$$

and $\frac{\partial w_1}{\partial l} = -\sin(\theta(l)) \neq 0$, the implicit function theorem states the existence of an open set $(\overline{u}_1-\epsilon, \overline{u}_1+\epsilon)$ containing $\overline{u}_1$ and a continuous function $\tilde{L}(u)$ defined on the open set such that

$$w_1(\tilde{L}(u), u) = \underline{u}_1, \quad \forall u \in (\overline{u}_1-\epsilon, \overline{u}_1+\epsilon).$$

Because the l that satisfies $w_1(l, u) = \underline{u}_1$ is unique, $L(u) = \tilde{L}(u)$ for all u in the neighborhood of $\overline{u}_1$. Therefore, L is continuous at $\overline{u}_1$.

Second, because $L(\overline{u}_1)$ is continuous in $\overline{u}_1$ and the composition of continuous functions is still continuous, $\Theta(\overline{u}_1) := \theta(L(\overline{u}_1), \overline{u}_1)$ is continuous in $\overline{u}_1$.

To simplify notation in Lemmas A.4-A.5, we express w_1 and w_2 as functions of θ. That is, $w_1(\theta)$ denotes $w_1(l(\theta))$, where $l(\theta)$ is the inverse of $\theta(l)$.

Lemma A.4. *Two curves w and $\tilde{w}$ start from initial conditions $(w_1(\theta^*), w_2(\theta^*), \theta^*)$ and $(\tilde{w}_1(\theta^*), \tilde{w}_2(\theta^*), \theta^*)$, respectively. Suppose $w_1(\theta^*) \geq \tilde{w}_1(\theta^*)$, $w_2(\theta^*) \leq \tilde{w}_2(\theta^*)$, $\boldsymbol{N}(\theta^*)w(\theta^*) = \boldsymbol{N}(\theta^*)\tilde{w}(\theta^*)$, w solves* (2.10)-(2.12)*, and $\tilde{w}$ satisfies* (2.10)-(2.11) *and*

$$\tilde{\kappa}(\tilde{\theta}) = \frac{d\tilde{\theta}}{dl} < \max_{x\in[0,1]} \frac{2\boldsymbol{N}(\tilde{\theta})((p_1(x), p_2(x)) - \tilde{w})}{r|\phi(\tilde{\theta})|^2}. \tag{A.6}$$

Then $\kappa(\theta) > \tilde{\kappa}(\theta)$, $w_1(\theta) > \tilde{w}_1(\theta)$, and $w_2(\theta) < \tilde{w}_2(\theta)$ for all $\theta > \theta^*$.

Proof: First, $\mathbf{N}(\theta^*)w(\theta^*) = \mathbf{N}(\theta^*)\tilde{w}(\theta^*)$ and (A.6) imply $\kappa(\theta^*) > \tilde{\kappa}(\theta^*)$. It follows from continuity that $\kappa(\theta) > \tilde{\kappa}(\theta)$ for θ near θ^*.

Second, we show that $\kappa(\theta) > \tilde{\kappa}(\theta)$ for all $\theta > \theta^*$. Suppose not, let θ^{**} be the first

$\theta > \theta^*$ such that $\kappa(\theta) = \tilde{\kappa}(\theta)$. Then

$$\begin{aligned}
& \cos(\theta^{**})w_1(\theta^{**}) + \sin(\theta^{**})w_2(\theta^{**}) \\
= \; & \cos(\theta^{**})\left(w_1(\theta^*) - \int_{\theta^*}^{\theta^{**}} \frac{\sin(\theta)}{\kappa(\theta)}d\theta\right) + \sin(\theta^{**})\left(w_2(\theta^*) + \int_{\theta^*}^{\theta^{**}} \frac{\cos(\theta)}{\kappa(\theta)}d\theta\right) \\
= \; & \cos(\theta^{**})w_1(\theta^*) + \sin(\theta^{**})w_2(\theta^*) + \int_{\theta^*}^{\theta^{**}} \frac{\sin(\theta^{**}-\theta)}{\kappa(\theta)}d\theta \\
< \; & \cos(\theta^{**})w_1(\theta^*) + \sin(\theta^{**})w_2(\theta^*) + \int_{\theta^*}^{\theta^{**}} \frac{\sin(\theta^{**}-\theta)}{\tilde{\kappa}(\theta)}d\theta \\
\leq \; & \cos(\theta^{**})\tilde{w}_1(\theta^*) + \sin(\theta^{**})\tilde{w}_2(\theta^*) + \int_{\theta^*}^{\theta^{**}} \frac{\sin(\theta^{**}-\theta)}{\tilde{\kappa}(\theta)}d\theta \\
= \; & \cos(\theta^{**})\tilde{w}_1(\theta^{**}) + \sin(\theta^{**})\tilde{w}_2(\theta^{**}),
\end{aligned}$$

where the first inequality follows from $\kappa(\theta) > \tilde{\kappa}(\theta), \forall\theta \in (\theta^*, \theta^{**})$, and the second inequality follows from

$$\begin{aligned}
& \cos(\theta^{**})(\tilde{w}_1(\theta^*) - w_1(\theta^*)) + \sin(\theta^{**})(\tilde{w}_2(\theta^*) - w_2(\theta^*)) \\
= \; & (\cos(\theta^{**}) - \cos(\theta^*))(\tilde{w}_1(\theta^*) - w_1(\theta^*)) + (\sin(\theta^{**}) - \sin(\theta^*))(\tilde{w}_2(\theta^*) - w_2(\theta^*)) \\
\geq \; & 0,
\end{aligned}$$

which further follows from $w_1(\theta^*) \geq \tilde{w}_1(\theta^*)$, $w_2(\theta^*) \leq \tilde{w}_2(\theta^*)$, $\cos(\theta^{**}) \leq \cos(\theta^*)$, and $\sin(\theta^{**}) \geq \sin(\theta^*)$. Therefore, (2.12) and (A.6) imply

$$\begin{aligned}
\kappa(\theta^{**}) &= \max_{x\in[0,1]} \frac{2\mathbf{N}(\theta^{**})((p_1(x), p_2(x)) - w(\theta^{**}))}{r|\phi(\theta^{**})|^2} \\
&> \max_{x\in[0,1]} \frac{2\mathbf{N}(\theta^{**})((p_1(x), p_2(x)) - \tilde{w}(\theta^{**}))}{r|\phi(\theta^{**})|^2} \\
&> \tilde{\kappa}(\theta^{**}),
\end{aligned}$$

which contradicts the definition of θ^{**}.

Third, $w_1(\theta) > \tilde{w}_1(\theta)$ and $w_2(\theta) < \tilde{w}_2(\theta)$ because

$$\begin{aligned} w_1(\theta) &= w_1(\theta^*) - \int_{\theta^*}^{\theta} \sin(x)/\kappa(x)dx > \tilde{w}_1(\theta^*) - \int_{\theta^*}^{\theta} \sin(x)/\tilde{\kappa}(x)dx = \tilde{w}_1(\theta), \\ w_2(\theta) &= w_2(\theta^*) + \int_{\theta^*}^{\theta} \cos(x)/\kappa(x)dx < \tilde{w}_2(\theta^*) + \int_{\theta^*}^{\theta} \cos(x)/\tilde{\kappa}(x)dx = \tilde{w}_2(\theta). \end{aligned}$$

Recall that if a curve starts from $(w_1, w_2) = (\overline{u}_1, \underline{u}_2)$, then $\Theta(\overline{u}_1)$ denotes the angle of the curve when it crosses **Y**. The following lemma shows an important property of Θ.

Lemma A.5. *If $u < \tilde{u}$ and $\Theta(u) = \Theta(\tilde{u})$, then $\Theta(\lambda u + (1-\lambda)\tilde{u}) > \Theta(u) = \Theta(\tilde{u})$ for $\lambda \in (0,1)$.*

Proof: Let w and $\tilde{w}$ be the solutions to (2.10)-(2.12) that start from $(\overline{u}_1 = u, \underline{u}_2)$ and $(\overline{u}_1 = \tilde{u}, \underline{u}_2)$, respectively. Construct a curve h as the convex combination of w and $\tilde{w}$:

$$h(\theta) = \lambda w(\theta) + (1-\lambda)\tilde{w}(\theta).$$

First, curve h satisfies

$$\kappa^h(\theta) < \max_{x\in[0,1]} \frac{2\mathbf{N}(\theta)((p_1(x), p_2(x)) - h(\theta))}{r|\phi(\theta)|^2}. \tag{A.7}$$

To prove (A.7), we show the inequality (A.8) below, which is equivalent to (A.7).

$$\begin{aligned} &\max_{x\in[0,1]} \frac{2\mathbf{N}(\theta)((p_1(x), p_2(x)) - h(\theta))}{r|\phi(\theta)|^2} \\ =\ & \frac{2\mathbf{N}(\theta)((p_1(x^*), p_2(x^*)) - h(\theta))}{r|\phi(\theta)|^2} \\ =\ & \lambda\frac{2\mathbf{N}(\theta)((p_1(x^*), p_2(x^*)) - w(\theta))}{r|\phi(\theta)|^2} + (1-\lambda)\frac{2\mathbf{N}(\theta)((p_1(x^*), p_2(x^*)) - \tilde{w}(\theta))}{r|\phi(\theta)|^2} \\ =\ & \lambda\kappa(\theta) + (1-\lambda)\tilde{\kappa}(\theta), \end{aligned}$$

where x^* is the optimal solution in (A.7). Because function $\frac{1}{x}$ is strictly convex in x,

$$\begin{aligned}\left(\max_{x\in[0,1]} \frac{2\mathbf{N}(\theta)((p_1(x),p_2(x))-h(\theta))}{r|\phi(\theta)|^2}\right)^{-1} &= \frac{1}{\lambda\kappa(\theta)+(1-\lambda)\tilde{\kappa}(\theta)} \\ &< \lambda\frac{1}{\kappa(\theta)}+(1-\lambda)\frac{1}{\tilde{\kappa}(\theta)} \\ &= \frac{1}{\kappa^h(\theta)}, \end{aligned} \tag{A.8}$$

where the last equality follows from the fact that h is a convex combination of w and $\tilde{w}$, and $\lambda\frac{dw}{d\theta}+(1-\lambda)\frac{d\tilde{w}}{d\theta}=\frac{dh}{d\theta}$.

Second, let $\hat{w}$ be the solution curve starting from $(\overline{u}_1=\lambda u+(1-\lambda)\tilde{u},\underline{u}_2)$. Lemma A.4 shows that $\hat{w}_1(\theta)>h_1(\theta)$ for all θ. In particular,

$$\begin{aligned}\hat{w}_1(\Theta(u))>h_1(\Theta(u)) &= \lambda w_1(\Theta(u))+(1-\lambda)\tilde{w}_1(\Theta(u)) \\ &= \lambda\underline{u}_1+(1-\lambda)\underline{u}_1=\underline{u}_1.\end{aligned}$$

Because the curve $\hat{w}$ has not reached **Y** at angle $\Theta(u)$, the angle at which $\hat{w}$ reaches **Y**, $\Theta(\lambda u+(1-\lambda)\tilde{u})$, is above $\Theta(u)$.

Proof of Theorem 1: First, we show that $\Theta(\overline{u}_1=\underline{u}_1)=\underline{\theta}$ and $\Theta(\overline{u}_1=\overline{\overline{u}}_1)=\underline{\theta}$. The former is because $L(\overline{u}_1=\underline{u}_1)=0$ and $\theta(l=0)=\underline{\theta}$. The latter follows from the fact that the solution curve starting from $(w_1,w_2,\theta)=(\overline{\overline{u}}_1,\underline{u}_2,\underline{\theta})$ is a straight line (see part (i) of Lemma A.2).

Second, define $\overline{u}_1^*$ as a maximizer of function Θ on $[\underline{u}_1,\overline{\overline{u}}_1]$. Such a maximizer exists because Θ is shown to be a continuous function in Lemma A.3.

Third, we show that Θ is strictly increasing in $[\underline{u}_1,\overline{u}_1^*]$; the proof that Θ is strictly decreasing in $[\overline{u}_1^*,\overline{\overline{u}}_1]$ is similar and hence omitted. By contradiction, suppose $\Theta(u^1)\geq\Theta(u^2)$ for some u^1 and u^2, where $0\leq u^1<u^2\leq\overline{u}_1^*$. Because Θ is continuous in

$[\overline{u}_1^*, \overline{\overline{u}}_1]$ and $\Theta(u^1) \in [\underline{\theta}, \Theta(\overline{u}_1^*)]$, the intermediate value theorem states the existence of $u^3 \in [\overline{u}_1^*, \overline{\overline{u}}_1]$ such that $\Theta(u^1) = \Theta(u^3)$. If $u^2 < \overline{u}_1^*$, then $u^2 < u^3$ and $u^2 \in (u^1, u^3)$. Lemma A.5 shows that $\Theta(u^2) > \Theta(u^1) = \Theta(u^3)$, which contradicts $\Theta(u^1) \geq \Theta(u^2)$. If $u^2 = \overline{u}_1^*$, then Lemma A.5 shows that $\Theta(\frac{u^1+u^3}{2}) > \Theta(u^1) \geq \Theta(u^2) = \Theta(\overline{u}_1^*)$. That $\Theta(\frac{u^1+u^3}{2}) > \Theta(\overline{u}_1^*)$ contradicts the fact that $\overline{u}_1^*$ is a maximizer.

Lemma A.6. *If $w_1 > \underline{u}_1$, $w_2 > \underline{u}_2$, and $w = (w_1, w_2) \in \partial\mathcal{E}$, then there is no payment at w.*

Proof: By contradiction, suppose there is payment at w and $(\tilde{w}_1, \tilde{w}_2)$ are the continuation payoffs after payment; then

$$\begin{aligned} w_1 + w_2 = \tilde{w}_1 + \tilde{w}_2 - rC &\leq \max_{(w_1,w_2)\in\mathcal{E}} w_1 + w_2 - rC \\ &= \overline{u}_1 + \underline{u}_2 \\ &= \underline{u}_1 + \overline{u}_2, \end{aligned}$$

where the last two equalities are (2.13) and (2.14). This implies that w is weakly below the line segment connecting $(\overline{u}_1, \underline{u}_2)$ and $(\underline{u}_1, \overline{u}_2)$. If $w_1 + w_2 < \underline{u}_1 + \overline{u}_2$, then w is strictly below the line, and hence in the interior of the triangle with vertices $(\underline{u}_1, \underline{u}_2)$, $(\overline{u}_1, \underline{u}_2)$, and $(\underline{u}_1, \overline{u}_2)$. Since the triangle is a subset of $\mathcal{E}$, w is also in the interior of $\mathcal{E}$, contradicting the assumption that $w \in \partial\mathcal{E}$. If $w_1 + w_2 = \underline{u}_1 + \overline{u}_2$, then define $\hat{w} := w + \epsilon(w - w^*)$, where $\epsilon > 0$ is a small number and w^* is a solution to $\max_{(w_1,w_2)\in\mathcal{E}} w_1 + w_2$. Because $\hat{w}$ is strictly below the line connecting $(\overline{u}_1, \underline{u}_2)$ and $(\underline{u}_1, \overline{u}_2)$, the above argument shows that $\hat{w}$ is in the interior of $\mathcal{E}$. Now $w = \frac{\epsilon}{1+\epsilon} w^* + \frac{1}{1+\epsilon}\hat{w}$ is a convex combination of w^* and an interior point $\hat{w}$, and therefore, w is interior too. This again contradicts the assumption that $w \in \partial\mathcal{E}$.

Lemma A.7. *If $\underline{u}_1 + \underline{u}_2 < \max_{(w_1,w_2)\in\mathcal{E}} w_1 + w_2 - rC$, then $\partial\mathcal{E}$ contains a horizontal portion, a vertical portion, and a downward sloping portion.*

Proof: We first characterize the horizontal portion and the vertical portion of $\partial\mathcal{E}$. Define $\overline{u}_i$ as in (2.5). The horizontal portion and vertical portion of $\partial\mathcal{E}$ are, respectively, $\{(w_1, \underline{u}_2) : \underline{u}_1 \leq w_1 \leq \overline{u}_1\}$ and $\{(\underline{u}_1, w_2) : \underline{u}_2 \leq w_2 \leq \overline{u}_2\}$. To see that these boundaries are non-degenerate, we need to show $\underline{u}_i < \overline{u}_i$. Equation (2.13) and the assumption $\underline{u}_1 + \underline{u}_2 < \max_{(w_1,w_2)\in\mathcal{E}} w_1 + w_2 - rC$ imply

$$\underline{u}_1 < \max_{(w_1,w_2)\in\mathcal{E}} w_1 + w_2 - rC - \underline{u}_2 = \overline{u}_1 + \underline{u}_2 - \underline{u}_2 = \overline{u}_1.$$

Similarly, we can show that $\underline{u}_2 < \overline{u}_2$.

Then, we study the portion of the boundary from $(\overline{u}_1, \underline{u}_2)$ to $(\underline{u}_1, \overline{u}_2)$. This portion stays above the straight line connecting the two points because $\mathcal{E}$ is convex. As we argue in the text, the portion satisfies the ODE in (2.10)-(2.12). Recall that part (ii) in Lemma A.2 shows that $\theta < \pi/2$ on any solution to the ODE in (2.10)-(2.12). Similar steps can show that $0 < \theta < \pi/2$ on any solution to the ODE. Therefore, the boundary from $(\overline{u}_1, \underline{u}_2)$ to $(\underline{u}_1, \overline{u}_2)$ is downward sloping.

Proof of Lemma 6: First, we show that $\overline{u}_1(w_1^*)$ is strictly increasing in w_1^* and $\overline{u}_2(w_1^*)$ is strictly decreasing in w_1^*. We only provide the proof for $\overline{u}_2(\cdot)$ as the proof for $\overline{u}_1(\cdot)$ is symmetric. For any two payoffs $w_1^* > \tilde{w}_1^*$, let w and $\tilde{w}$ be the curves starting at $(w_1^*, S - w_1^*, \frac{\pi}{4})$ and $(\tilde{w}_1^*, S - \tilde{w}_1^*, \frac{\pi}{4})$, respectively. The proof of $\overline{u}_2(w_1^*) < \overline{u}_2(\tilde{w}_1^*)$ consists of parts (i), (ii), and (iii). During the proof, we shall write w_2, θ, and κ as functions of w_1.

1. If $w_1 = \tilde{w}_1^*$, then $\theta(w_1) > \pi/4 = \tilde{\theta}(w_1)$. Furthermore, $w_2(\tilde{w}_1^*) < \tilde{w}_2(\tilde{w}_1^*)$ because

$$\begin{aligned} w_2(\tilde{w}_1^*) &= w_2(w_1^*) + \int_{w_1^*}^{\tilde{w}_1^*} w_2'(w_1)dw_1 \\ &= w_2(w_1^*) + \int_{\tilde{w}_1^*}^{w_1^*} \cot(\theta(w_1))dw_1 \\ &< w_2(w_1^*) + \int_{\tilde{w}_1^*}^{w_1^*} 1dw_1 \\ &= S - \tilde{w}_1^* = \tilde{w}_2(\tilde{w}_1^*). \end{aligned}$$

2. If there is some $\overline{w}_1$ such that $\theta(w_1) > \tilde{\theta}(w_1)$ for all $w_1 \in (\overline{w}_1, \tilde{w}_1^*]$, then $w_2(\overline{w}_1) < \tilde{w}_2(\overline{w}_1)$. The proof given below is similar to that in part (i).

$$\begin{aligned} w_2(\overline{w}_1) &= w_2(\tilde{w}_1^*) + \int_{\tilde{w}_1^*}^{\overline{w}_1} w_2'(w_1)dw_1 \\ &= w_2(\tilde{w}_1^*) + \int_{\overline{w}_1}^{\tilde{w}_1^*} \cot(\theta(w_1))dw_1 \\ &< \tilde{w}_2(\tilde{w}_1^*) + \int_{\overline{w}_1}^{\tilde{w}_1^*} \cot(\tilde{\theta}(w_1))dw_1 \\ &= \tilde{w}_2(\overline{w}_1), \end{aligned}$$

   where the inequality follows from $w_2(\tilde{w}_1^*) < \tilde{w}_2(\tilde{w}_1^*)$ in part (i) and the assumption that $\theta(w_1) > \tilde{\theta}(w_1)$ for all $w_1 \in (\overline{w}_1, \tilde{w}_1^*]$.

3. We show $\theta(w_1) > \tilde{\theta}(w_1)$ for all $w_1 \in [\underline{u}_1, \tilde{w}_1^*]$ by contradiction. Suppose not, let $\overline{w}_1$ be the largest $w_1 < \tilde{w}_1^*$ such that $\theta(w_1) = \tilde{\theta}(w_1)$. That is, $\theta(w_1) > \tilde{\theta}(w_1)$ for all $w_1 \in (\overline{w}_1, \tilde{w}_1^*]$ and $\theta(w_1) = \tilde{\theta}(w_1)$ for $w_1 = \overline{w}_1$. Then part (ii) implies $w_2(\overline{w}_1) < \tilde{w}_2(\overline{w}_1)$. Therefore, $\mathbf{N}(\theta(\overline{w}_1))(\overline{w}_1, w_2(\overline{w}_1)) < \mathbf{N}(\tilde{\theta}(\overline{w}_1))(\overline{w}_1, \tilde{w}_2(\overline{w}_1))$ and $\kappa(\overline{w}_1) > \tilde{\kappa}(\overline{w}_1)$.

   On the other hand, $\frac{d\theta}{dw_1} \geq \frac{d\tilde{\theta}}{dw_1}$ at $w_1 = \overline{w}_1$ because $\theta(w_1) > \tilde{\theta}(w_1)$ for all $w_1 \in$

$(\overline{w}_1, \tilde{w}_1^*]$ and $\theta(w_1) = \tilde{\theta}(w_1)$ for $w_1 = \overline{w}_1$. Therefore,

$$\kappa(\overline{w}_1) = \frac{d\theta}{dl} = \frac{d\theta}{dw_1}\frac{dw_1}{dl} = -\frac{d\theta}{dw_1}\sin(\theta(\overline{w}_1)) \leq -\frac{d\tilde{\theta}}{dw_1}\sin(\tilde{\theta}(\overline{w}_1)) = \tilde{\kappa}(\overline{w}_1),$$

which contradicts the inequality $\kappa(\overline{w}_1) > \tilde{\kappa}(\overline{w}_1)$ shown above. Therefore, $\theta(w_1) > \tilde{\theta}(w_1)$ for all $w_1 \in [\underline{u}_1, \tilde{w}_1^*]$ and by part (ii), $w_2(\underline{u}_1) < \tilde{w}_2(\underline{u}_1)$. This finishes the proof because $\overline{u}_2(w_1^*) = w_2(\underline{u}_1)$ and $\overline{u}_2(\tilde{w}_1^*) = \tilde{w}_2(\underline{u}_1)$.

Second, we show the existence and the uniqueness of w_1^* that satisfies $\underline{u}_1 + \overline{u}_2(w_1^*) = \overline{u}_1(w_1^*) + \underline{u}_2$. It follows from

$$\lim_{w_1^* \downarrow \underline{u}_1} \overline{u}_2(w_1^*) = S - \underline{u}_1, \qquad \lim_{w_1^* \uparrow S - \underline{u}_2} \overline{u}_1(w_1^*) = S - \underline{u}_2$$

that

$$\begin{aligned}
\lim_{w_1^* \downarrow \underline{u}_1} \overline{u}_1(w_1^*) + \underline{u}_2 \quad &< \quad (S - \underline{u}_2) + \underline{u}_2 = \lim_{w_1^* \downarrow \underline{u}_1} \overline{u}_2(w_1^*) + \underline{u}_1, \\
\lim_{w_1^* \uparrow S - \underline{u}_2} \overline{u}_1(w_1^*) + \underline{u}_2 \quad &= \quad (S - \underline{u}_2) + \underline{u}_2 > \lim_{w_1^* \uparrow S - \underline{u}_2} \overline{u}_2(w_1^*) + \underline{u}_1.
\end{aligned}$$

The above inequalities and the intermediate value theorem imply the existence of a w_1^* satisfying $\underline{u}_1 + \overline{u}_2(w_1^*) = \overline{u}_1(w_1^*) + \underline{u}_2$. The uniqueness follows from the monotonicity of $\overline{u}_1(w_1^*) - \overline{u}_2(w_1^*)$ in w_1^*, and the monotonicity is shown in the first step.

Proof of Theorem 2: We can express equation (2.17) equivalently as

$$1 + B(D-S)S - \sqrt{1 + 2B(D-S)S} = rCB(D-S). \tag{A.9}$$

The rest of this proof consists of five steps.

First, $g(S) := 1 + B(D-S)S - \sqrt{1 + 2B(D-S)S}$ is symmetric around $D/2$. It is increasing on $[0, D/2]$ because

$$g'(S) = \left(1 - \frac{1}{\sqrt{1 + 2B(D-S)S}}\right) B(D - 2S) > 0, \quad \forall S \in (0, D/2).$$

Second, there is a cutoff $S^* \in (0, D/2)$ such that g is convex in $(0, S^*)$ and concave in $(S^*, D/2)$. The second derivative of g is

$$\begin{aligned} g''(S) &= -2B\left(1 - \frac{1}{\sqrt{1 + 2B(D-S)S}}\right) + \frac{1}{(\sqrt{1 + 2B(D-S)S})^3} B^2(D - 2S)^2 \\ &= \frac{B^2(D-2S)^2 - 2B(1 + 2B(D-S)S)(\sqrt{1 + 2B(D-S)S} - 1)}{(\sqrt{1 + 2B(D-S)S})^3}. \end{aligned} \quad \text{(A.10)}$$

It is easy to verify that $g''(S = 0) > 0$ and $g''(S = D/2) < 0$. Moreover, the numerator in (A.10) is monotonically decreasing in $S \in (0, D/2)$. Therefore, there is a cutoff $S^* \in (0, D/2)$ such that

$$g''(S) \begin{cases} > 0, & \text{if } S \in (0, S^*); \\ < 0, & \text{if } S \in (S^*, D - S^*); \\ > 0, & \text{if } S \in (D - S^*, D). \end{cases}$$

Third, generically, (A.9) has two or more solutions in $(0, D)$ if it has at least one solution.[8] Suppose $\overline{S}$ is the largest solution; then there should be another solution below $\overline{S}$. To see this, note that g is below $rCB(D - S)$ when S is close to D, since $g'(S = D) = 0$. That means g stays below the straight line $rCB(D-S)$ for all $S \in (\overline{S}, D)$ and stays above the straight line for S slightly below $\overline{S}$. But $g(S = 0) = 0$ is below the straight line, which means there is another solution below $\overline{S}$.

Fourth, if $\overline{S} \in (D - S^*, D)$ is a solution, then $g(S) > rCB(D - S)$ for all $S \in$

[8]We will ignore the case where (A.9) has no solution. In this case, C is so large that no feasible contract exists except autarky.

$[D - S^*, \overline{S})$. This follows from the convexity of g on $[D - S^*, D]$ and the fact that D is also a solution. In particular, g is above the straight line at $S = D - S^*$.

Fifth, suppose $\underline{S}$ and $\overline{S}$ are, respectively, the smallest and the largest solutions in $(0, D)$. We show that no other solution exists between $\underline{S}$ and $\overline{S}$. There are three possibilities:

1. $\underline{S} \in (0, S^*)$. Then $\overline{S} > D - S^*$ because $g(\underline{S}) > g(\overline{S})$. The fourth step and $\overline{S} \in (D - S^*, D)$ imply that $g(S)$ is above the straight line for all $S \in (D - S^*, \overline{S})$. $g(S)$ is above the straight line for all $S \in (\underline{S}, S^*]$ because g is increasing but the straight line is decreasing. $g(S)$ is above the straight line for all $S \in [S^*, D - S^*]$ because g is concave and both $g(S^*)$ and $g(D - S^*)$ are above the straight line.

2. $\underline{S} \in [S^*, D - S^*]$. If $\overline{S} \in [S^*, D - S^*]$, then the conclusion follows from the concavity of g. If $\overline{S} \in (D - S^*, D)$, then the proof is similar to that in part (i).

3. $\underline{S} \in (D - S^*, D)$. Then $\overline{S} \in (D - S^*, D)$. But the fourth step implies $g(S) > rCB(D - S)$ for all $S \in [D - S^*, \overline{S})$. This contradicts the assumption that $\underline{S} \in (D - S^*, D)$ is a solution.

APPENDIX B

APPENDIX OF SECTION THREE

Proof of Proposition 1

In any bubbly equilibrium

$$\begin{aligned}\frac{Q^b_{t+1}H_{t+1}}{Q^b_tH_t} &= \frac{\phi_{t+1}R_{t+1}K_{t+1}}{\phi_tR_tK_t},\\ \frac{R_{t+1}g_t}{(1-\delta)p} &= \frac{\phi_{t+1}R_{t+1}K_{t+1}}{\phi_t\frac{K_{t+1}(1-\psi)\alpha}{(1-\phi_t)\psi}},\\ \frac{\phi_{t+1}}{\phi_t} &= \frac{(1-\psi)\alpha g_t}{(1-\delta)\psi(1-\phi_t)p} = \frac{zg_t}{(1-\phi_t)},\end{aligned} \tag{B.1}$$

where $z \equiv \frac{(1-\psi)\alpha}{(1-\delta)\psi p}$. And it is trivial that in a bubbly steady state, $\phi^* = 1 - z$. Equation (B.1) is important to understand the dynamics of bubble. The left part of this chapter is how we rely on Equation (B.1) to prove Proposition 1, and it includes the following six lemmas.

Lemma B.1. *In any equilibrium, $\{\phi_t\}$ converge to either bubbly steady state or bubbleless steady state.*

Proof: First, we argue that the sequence $\{\phi_t\}$ has a limitation. We prove it by contradiction. Assume there is a sequence $\{\phi_t\}$ which is not convergent. Define two sequences $\{\overline{g}_t\} \equiv sup(\{g_{t+i}\}_{i=0}^{i=\infty})$ and $\{\underline{g}_t\} \equiv inf(\{g_{t+i}\}_{i=0}^{i=\infty})$. At time t, if $\phi_t > 1 - z\underline{g}_t$, then $\frac{z\underline{g}_t}{1-\phi_t} > 1$ while $\phi_{t+i} \geq [\frac{z\underline{g}_t}{1-\phi_t}]^i\phi_t$ for any $i > 0$. The sequence will finally excess 1. And if $\phi_t < 1 - z\overline{g}_t$, then $\frac{z\overline{g}_t}{1-\phi_t} < 1$ while $\phi_{t+i} \leq [\frac{z\overline{g}_t}{1-\phi_t}]^i\phi_t$ for any $i > 0$. Finally, the sequence will converge to zero. Thus $\phi_t \in [1 - z\overline{g}_t, 1 - z\underline{g}_t]$. Since this should be held for any t, $lim_{t\to\infty}[1 - z\overline{g}_t] \leq lim_{t\to\infty}\phi_t \leq lim_{t\to\infty}[1 - z\underline{g}_t]$, and $lim_{t\to\infty}\phi_t = 1 - z$. It is contradicted with the assumption $\{\phi_t\}$ which is not convergent. By this analysis, we can also find $\{\phi_t\}$ will converge to either bubbly steady state $\hat{\phi}^* = 1 - z$ or bubbleless steady state 0.

Lemma B.2. *If $\phi_t > \phi'_t$, then $\phi_{t+i} > \phi'_{t+i}$ for any $i > 0$.*

Proof: If $\phi_t > \phi'_t$, by equation (B.1), we have $\phi_{t+1} = zg_t(\frac{1}{1-\phi_t} - 1) > zg_t(\frac{1}{1-\phi'_t} - 1) = \phi'_t$. By induction, we know $\phi_{t+i} > \phi'_{t+i}$ for any $i > 0$.

Lemma B.3. *There is an nonempty, open, convex set Φ_0 containing all ϕ_0, from which there are corresponding asymptotically bubbleless equilibrium.*

Proof: To prove this lemma, we have three steps.

First we prove non-emptiness by arguing that when ϕ_0 is sufficient small, there is always a asymptotically bubbleless equilibrium. If $\{g_t\}$ is an increasing sequence, then $g_t < 1$. Choose $\phi_0 < 1 - z$, then $\frac{\phi_{t+1}}{\phi_t} < \frac{z}{1-\phi_0} < 1$. Thus $\lim_{t\to\infty} \phi_t = 0$. If $\{g_t\}$ is a decreasing sequence, since $\{g_t\}$ converges to 1, there is a time T such that $zg_T < 1$. Choose $\phi_0 < min[\frac{1}{2(2zg_0)^T}, \frac{1-zg_T}{(2zg_0)^T)}]$ satisfying that $\phi_t < \frac{1}{2}$ when $t \in (0, T]$ and $\phi_T < 1 - zg_T$. Then $\frac{\phi_{T+i+1}}{\phi_{T+i}} < \frac{zg_T}{\phi_T} < 1$. Then $\{\phi_t\}$ converges to zero. And we can construct a asymptotically bubbleless equilibrium with this $\{\phi_t\}$.

Second, we prove the set is open. In step 1, we already show the set is open to the left. Now, we show the set is open to the right. We argue that if $\phi_0 \in \Phi_0$ then $\phi_0 + \eta \in \Phi_0$ when η is sufficient small. First we consider $\{g_t\}$ is increasing sequence. Since $\lim_{t\to\infty} \phi_t = 0$, there is a time T when $\phi_T < 1 - z$. Choose $\eta = \frac{1-z-\phi_T}{2(\frac{z}{1-\bar{\phi}})^T}$ where $\bar{\phi} = max[\{\phi_t\}_0^T]$, and define a new sequence $\{\phi'_t\}$ starting at $\phi_0 + \eta$. $\phi'_T < \phi_T + \frac{1-z-\phi_T}{2} < 1 - z$ then any ϕ'_{T+i} will be smaller then $1 - z$, and $\{\phi'_t\}$ converges to zero.If $\{g_t\}$ is a decreasing sequence, since $\{\frac{zg_t}{1-\phi_t}\}$ converges to z, there is a time T after when $\frac{zg_t}{1-\phi_t} < 1$ for all t. Choose

$\eta = \frac{1-zg_T-\phi_T}{2(\frac{zg_0}{1-\bar{\phi}})^T}$ where $\bar{\phi}$ is defined the same as before. Starting from $\phi_t + \eta$ there is a new sequence named as $\{\phi'_t\}$. $\phi'_T < \phi_T + \frac{1-zg_T-\phi_T}{2} < 1 - zg_T$ and $\frac{zg_t}{1-\phi'_t} < \frac{zg_T}{1-\phi'_T}$ for all $t \geq T$. Then sequence $\{\phi'_t\}$ converges to 0.

Third, we prove convexity. Assume ϕ_0 and ϕ'_0 ($\phi_0 > \phi'_0$) are in Φ_0, then for any $\phi''_0 \in (\phi'_0, \phi_0)$ by lemma 2 $\phi_t > \phi''_t > \phi'_t$. We have $0 = \lim_{t\to\infty} \phi_t \geq \lim_{t\to\infty} \phi''_t \geq \lim_{t\to\infty} \phi'_t = 0$. Thus, $\phi''_0 \in \Phi_0$.

Lemma B.4. *There is a $\hat{\phi}_0$ as the supremum of set Φ_0. Starting from $\hat{\phi}_0$, there is an corresponding equilibrium.*

Proof: Since set Φ_0 is a bounded open set, then there exists a supremum denoted as $\hat{\phi}_0$ We claim there exists an equilibrium starting at $\hat{\phi}_0$. Define a sequence $\{\phi_t\}$ starting at $\hat{\phi}_0$ satisfying equation (B.1). If there is no equilibrium starting from $\hat{\phi}_0$, then there must be some time T when $\phi_T \geq 1$. Now we claim the sequence $\{\phi'_t\}$ starting at $\hat{\phi}_0 - \eta$ also corresponds to no equilibrium when η is sufficient small. If $\phi_T > 1$, then choose $\eta = \frac{\phi_T - 1}{2\Delta^T}$ where $\Delta \equiv \max(\{\frac{zg_t}{1-\phi_t}\}_0^{T-1})$. Then $\phi'_T \geq \phi_T - \frac{\phi_T-1}{2} > 1$. If $\phi_T = 1$, choose $\eta = \frac{zg_T}{2(zg_T+1)\Delta^T}$. Then $\phi'_T > 1 - \frac{zg_T}{2(zg_T+1)}$, which makes $\phi'_{T+1} = zg_T \frac{\phi'_T}{1-\phi'_T} \geq 2 + zg_T > 1$.Since there is no corresponding equilibrium starting from $\hat{\phi} - \eta$, then $\hat{\phi}_0 - \eta \notin \Phi_0$. At the same time, $\hat{\phi}_0 - \eta > \phi_0$ for any $\phi_0 \in \Phi_0$ because Φ_0 is a convex set. It is contradicted that $\hat{\phi}_0$ is the supremum of set Φ_0.

Lemma B5. *There is a unique value of ϕ_0, which corresponds to a bubbly equilibrium.*

Proof: Since Lemma B.1, Lemma B.4 and Lemma B.5 imply that $\hat{\phi}_0$ corresponds to a bubbly equilibrium. Here we just prove the uniqueness. Assume that there are two initial values ϕ_0 and ϕ'_0 starting from which $\{\phi_t\}$ and $\{\phi'_t\}$ converges to the bubbly steady state. Since $\lim_{t\to\infty} \frac{zg_t}{1-\phi_t} = 1$ and $\lim_{t\to\infty}(1-\phi'_t) = z < 1$, then there exist a time T after when

$\frac{zg_t}{1-\phi_t} > 1 - \phi'_t$ for any $t \geq T$. We have $\phi_{T+1} - \phi'_{T+1} = \frac{zg_T}{(1-\phi_t)(1-\phi'_t)}(\phi_T - \phi'_T) > \phi_T - \phi'_T$. By induction, we know that $\phi_{T+i} - \phi'_{T+i} > \phi_T - \phi'_T$ for any $i > 0$. Then ϕ_{T+i} and ϕ'_{T+i} at least one are not in $(\hat{\phi}^* + \frac{|\phi_T - \phi'_T|}{2}, \hat{\phi}^* - \frac{|\phi_T - \phi'_T|}{2})$, which is contradicted with that both $\{\phi_t\}$ and $\{\phi'_t\}$ converge to $\hat{\phi}^*$.

Lemma B.6. *Starting from $\phi_0 > \hat{\phi}_0$, there is no equilibrium.*

Proof: By Lemma B.1, in any equilibrium ,starting from $\phi_0 > \hat{\phi}_0$, the sequence $\{\phi_t\}$ can only converges to 0 or $\hat{\phi}^*$. By Lemma B.5, starting from $\phi_0 > \hat{\phi}_0$, the sequence $\{\phi_t\}$ can not converge to $\hat{\phi}^*$. By Lemma B.2, the limitation of $\{\phi_t\}$ should be greater than or equal to $\hat{\phi}^*$, thus $\{\phi_t\}$ can not converge to zero. So, starting from $\phi_0 > \hat{\phi}_0$, there is no equilibrium.