Local Direct Elections, Local Information, and Meritocratic Selection*

Yu Wang† and Renfu Luo‡

Job Market Paper
(November 25, 2019)

Click here for the current version and progress

Abstract

This paper studies the relationship between local direct elections and meritocratic selection through the mechanism of local information. This paper’s theoretical model based on the Bayesian inference framework shows that local direct elections in rural China facilitate the meritocratic selection of both village committee members and village party secretaries. Local direct elections transfer the authority of selecting village committee members, from township officials to village residents. Because village residents, compared to township officials, have advantages in the local information on village committee candidates, they infer those candidates’ virtue and capacity more accurately and precisely. The introduction of local direct elections, with such improved effectiveness of inference, enhances the expectations of the composite virtue-and-capacity of village committee members, while reduces the variances of those. Further, part of or all village committee members are also the candidates for village party secretaries. Therefore, with such improved candidate pools, the expectations of the composite virtue-and-capacity of village party secretaries are also enhanced.

Keywords: Local Direct Elections, Meritocratic Selection, Local Information

---

* We would like to thank Marc Bellemare, Loren Brandt, Jay Coggins, Paul Glewwe, John Freeman, Elton Mykerezi, Terry Roe, and Sean Sylvia for their valuable comments and suggestions. All errors are ours.
† Wang: Job market candidate. Department of Applied Economics, University of Minnesota (e-mail: wang5979@umn.edu).
‡ Luo: China Center for Agricultural Policies, School of Advanced Agricultural Sciences, Peking University (e-mail: luorf.ccap@pku.edu.cn).
When the Grand course was pursued, a public and common spirit ruled all under the sky; they chose men of talents, virtue, and ability; their words were sincere, and what they cultivated was harmony.” – Confucius (450 BC, translated by James Legge [1885]), Li Chi: Book of Rites.

“The aim of every political constitution, is or ought to be, first to obtain for rulers men who possess most wisdom to discern, and most virtue to pursue, the common good of society.” – Hamilton, Madison and Jay (1788 [2008]), The Federalist Papers.

I. Introduction

Meritocratic selection is pursued around the world, both in ancient and modern times. Dating back to around 500 BC, Chinese politicians and philosophers argued that those who govern should be selected by merit rather than inherited status (Sienkewicz, 2003). When the concept of meritocracy spread to Europe and the U.S., it was favored by philosophers (Kazin et al., 2010) and claimed in political statements (Hamilton, Maddison and Jay, 2008).

China has developed a series of top-down political selection schemes, which emphasize the assessment, recommendation, and promotion of politicians based on their virtue and capacity, aiming at ex ante meritocratic selection1. However, those top-down political selection schemes may suffer from adverse selection, thus their effects in the meritocratic selection of politicians could be limited.

Unlike China’s top-down scheme, an electoral system based on a bottom-up political selection scheme was established in ancient Western regimes and gradually prevailed2. However, most previous studies emphasize elections’ role in addressing moral hazard, that is, facilitating the ex post accountability of politicians (Laffont, 2000; Besley, 2005).

---

1 Around 134 BC, the assessment and recommendation system of noble families was established (Qian, 2012). By expanding enfranchisement, the civil examination system of scholars was developed around 605 AD, prevailing for more than 1,000 years and greatly influencing the political selection schemes of China and other countries (Elman, 2013; Bai and Jia, 2016; Bell, 2016).

2 Around 508 BC, Athenian democracy was established. By expanding enfranchisement, this electoral system has evolved into a modern representative democracy (Loeper, 2017).
This paper studies how local direct elections, by providing local information, address the adverse selection and facilitate the meritocratic selection of politicians. The introduction of local direct elections to Chinese local governance provides an institutional comparison to identify the relationship and mechanism between local direct elections and meritocratic selection. In addition, the Chinese local governance is a typical political selection context characterized as the small groups in the grassroots level of the stratified governance structure. Specifically, after introducing local direct elections, village leaders (the small group leaders) and other village committee members\(^3\) were no longer appointed by township officials, but directly elected by village residents. We build up a theoretical model, demonstrating that local direct elections facilitate the meritocratic selection of village committee members, by providing more local information on the virtue and capacity of village committee candidates. Village party secretaries, superior to village leaders in the governance ladders, remain being appointed by township officials. Even so, our theoretical model shows that local direct elections for village committee members facilitate the meritocratic selection of village party secretaries, because part of or all village committee members are also the candidates for village party secretaries.

Our model emphasizes that the essential mechanism through which local direct elections work in the meritocratic selection of politicians is the advantages in local information on the political candidates that local direct elections have; to frame this, our model uses the Bayesian inference framework. This is different from previous studies focusing the strategic behaviors between politicians and voters,\(^4\) and using the game theory framework. The introduction

\(^3\) In the following paragraphs, village committee members include village leaders who are chairs of village committees and the representative village committee members, as well as other village committee members.

of local direct elections transfers the authority of selecting village committee members from township officials to village residents. Village residents naturally communicate with village committee candidates more often than township officials, implying village residents’ advantage in obtaining local information about these candidates (Ghatak, 1999; Bell, 2016). Therefore, village residents, compared to township officials, can infer the virtue and capacity of village committee candidates more accurately and precisely; in other words, the effectiveness of inference is improved by local direct elections.

Due to village residents’ advantages in local information, local direct elections that empower them address the adverse selection, that is, facilitate the meritocratic selection of village committees. With the more accurate inference, our theoretical model proves that in a representative village, the expectation of the composite virtue-and-capacity (a weighted average of virtue and capacity) of each elected village committee member is greater than that of each appointed village committee member. Our model also proves, with the more precise inference, that the variance of the composite virtue-and-capacity of each elected village committee member is lower than that of each appointed village committee member. These theoretical findings of improved effectiveness of inference are in line with Hayek (1945) and Chan (2013) that assessment and decisions must be left to people with local information advantages.

The advantages in local information on political candidates aggregated by local direct elections also endow the performance-based promotion for superior politicians in a stratified governance structure with meritocratic selection. Our model further shows that local direct elections, by improving the candidate pools of village party secretaries, facilitate the meritocratic selection of village party secretaries, chairs of village party committees that oversee village committees. Part of or all village committee members, including village leaders,

5 “More accurately” means in the Bayesian inference, the posterior mean of virtue (or capacity) is closer to the real value of virtue (or capacity). “More precisely” means in the Bayesian inference, the posterior variance of virtue (or capacity) gets smaller. These will be discussed in Section III in details.
are also village party committee members and are therefore candidates for village party secretaries (O’Brien and Li, 2000). In a representative village, the local direct election increases the expectation of the composite virtue-and-capacity of each village committee member. In other words, the candidate pool for the village party secretary is improved. Therefore, our model shows that the expectation of the composite virtue-and-capacity of the village party secretary also increases. These theoretical findings of improved candidate pools imply that the local information provided by local direct elections directly benefit the bottom-up local political selection, but also indirectly benefits the top-down political selection in higher governance ladders.

The rest of this paper is organized as follows. Section II introduces the institutional background. Section III develops the theory. Finally, Section IV concludes this paper.

II. Local Governance in Rural China

The administrative organizations of Chinese villages consist of two committees. Village committees, de facto government entities at the village level, are chaired by village leaders and composed of other members. Village party committees, which represent the village-level leadership of the Chinese Communist Party (CCP), are chaired by village party secretaries and composed of other members. In practice, part of or all village committee members, especially village leaders, are also members of village party committees. Likewise, part of or all village party committee members are also members of village committees.

Village party committees oversee village committees, thus village party secretaries are superior to village leaders in the governance hierarchy, stipulated by the “Organic Law of Village Committees (OLVC)” (National People's Congress of China, 1998) and the “Working Regulation for Rural Grassroots Organizations of the Chinese Communist Party (WRRGOCCP)” (Central Committee of the Chinese Communist Party, 1999). Village committees are
responsible for providing village infrastructure and public services, developing the local economy, and improving village residents’ income (National People’s Congress of China, 1998; Martinez-Bravo et al., 2011). The role of village party committees in the development of the local economy is to approve village committees’ plans and monitor their implementation (Oi and Rozelle, 2000).

The likelihood of being selected either as a village committee member or as a village party committee member is positively associated with the candidate’s virtue and capacity (Bell, 2016; Tang, 2016), rooted in the concept and practice of meritocratic selection in Chinese history (Zhang, 2012). The selection of village committee members, including village leaders, requires that candidates be law-abiding, of moral integrity, intrinsically motivated to serve the village residents, and have a diploma and administrative capacity (National People’s Congress of China, 1998). The selection of village party committee members, including village party secretaries, requires candidates to have professional knowledge and skills, to be responsive to the needs and demands of the village residents, and to be intrinsically motivated to serve them (Central Committee of the Chinese Communist Party, 1999).

The selection scheme for village leaders and other village committee members has changed, from being appointed by the township officials to being directly elected by the village residents through local direct elections. The introduction of local direct elections in rural China has been a gradual process (Martinez-Bravo et al., 2014). By 2010, most villages in rural China have introduced local direct elections for village leaders and other village committee members (Padró i Miquel et al., 2015; Wong et al., 2017).

---

6 In 1998, the National People’s Congress of China passed the revised “Organic Law of Village Committees (OLVC),” introducing local direct elections for village leaders and other village committee members in rural China, resulting in the election of village leaders and other village committee members through open nomination and competitive elections (O’Brien and Han, 2009). After the national legislation in 1998, each province in China introduced its own “Provincial Measures for Implementing the Organic Law of Village Committees” to provide additional instructions on the implementation of local direct elections (O’Brien and Zhao, 2011). Counties and townships then followed (Wong et al., 2017).
In contrast, the selection scheme for village party secretaries and other village party committee members remains the responsibility of the township officials, by appointment (Central Committee of the Chinese Communist Party, 1999). Common village party committee members are elected by all village party members and then approved by the township officials. In other words, they have the final say on village party committee members. However, village party secretaries are either directly appointed or nominated by the township officials, and then elected by all village party members. In other words, the township officials have decisive authority over village party secretaries. Because village party secretaries have the highest authority within villages, they are usually promoted from village party committee members by the township officials, based on their performance in local governance.

III. Theory

Our theoretical model based on the Bayesian inference framework discusses how the introduction of local direct elections to a representative village facilitates the meritocratic selection of both village committee members and the village party secretary, by providing local information.

A. Inferences on Village Committee Candidates

In this section, we use the Bayesian inference framework to discuss the assumption about the virtue and capacity of village committee candidates in a representative village. Because the village leader is the chair of the village committee, we only discuss the village leader and consider her the representative of all village committee members. Thus, our theoretical findings in this section apply to other village committee members. We find that because the representative village resident has an advantage in terms of local information about village leader candidates, that is, she naturally communicates with the village leader candidates more often than the representative township official, her inferences of the virtue and capacity of these candidates are more
accurate and precise, namely improved *effectiveness of inference*.

1. **Setup:** Our theory discusses a representative village, in which all adult village residents are potential village leader candidates. Each one has two personal characteristics: *virtue* and *capacity*, both assumed to be independent and identically distributed on $[0,1]$ with a mean of 0.5.\(^7\)

After the introduction of local direct elections in this representative village, a pool of village leader candidates, a subset of all potential candidates, competes to be elected village leader by the village residents. Before this introduction, a pool of village leader candidates competed to be appointed village leader by the township officials. The virtue of village leader candidate \(i\) is denoted by \(\alpha_i\), with \(\alpha_i \in [0,1]\), and her capacity is denoted by \(\theta_i\), with \(\theta_i \in [0,1]\), where \(i = \{1,2,\ldots\}\). In the following analysis, we discuss one representative village resident or township official instead of village residents or township officials, assuming that village residents or township officials have homogenous perceptions in the inferences of the candidates of each village committee member, including the village leader.\(^9\)

2. **Bayesian Inferences:** The representative village resident or township official cannot observe each village leader candidate’s virtue or capacity. To

---

\(^7\) Virtue refers to characteristics including, but not limited to, law-abiding, of moral integrity, and intrinsically motivated to serve the village residents (Central Committee of the Chinese Communist Party, 1999; Dal Bó et al., 2017; National People’s Congress of China, 1998). Capacity refers to characteristics including, but not limited to, professional knowledge and administrative skills (Alesina and Tabellini, 2007; Central Committee of the Chinese Communist Party, 1999; Dal Bó et al., 2017; National People’s Congress of China, 1998).

\(^8\) The virtue and capacity of village residents are both assumed to be bounded, because (1) the number of residents in a village, a local area, is usually small (Liu et al., 2009; Martínez-Bravo et al., 2014), and (2) their socio-economic characteristics, thinking patterns, and behaviors tend to be homogenous, due to homogenous socio-economic, cultural, and institutional constraints, and long-term, generation-by-generation interactions in local areas (Bell, 2016). In a word, the virtue and capacity are local information. For simplicity, we assume that their virtue and capacity are both bounded at \([0,1]\).

\(^9\) The perceptions are both assumed homogenous, because, similar to before, (1) small number of village residents or township officials, and (2) Village residents or township officials have homogenous socio-economic, cultural, and institutional constraints.
infer the virtue and capacity of the village leader candidates, the representative village resident or township official, by naturally communicating with each candidate, obtains $\Omega^\alpha_{it}$, a series of observations of the virtue of village leader candidate $i$ at the natural communication of the $t$-th time, where $t = 1, ..., T^{Ele}$ and $1, ..., T^{App}$, respectively. $T^{Ele}$ or $T^{App}$ represents the ultimate times of natural communication between the representative village resident or township official and each village leader candidate, immediately before deciding which candidate to select as village leader. $\Omega^\alpha_{it}$ is given by

$$\Omega^\alpha_{it} = \alpha_i + v_{it}$$

where $v_{it}$ is a series of random shocks when observing virtue. Similarly, the representative village resident or township official, by naturally communicating with each village leader candidate, obtains $\Omega^\theta_{it}$, a series of observations of the capacity of village leader candidate $i$ at time $t = 1, ..., T^{Ele}$ and $1, ..., T^{App}$, respectively. $\Omega^\theta_{it}$ is given by

$$\Omega^\theta_{it} = \theta_i + \omega_{it}$$

where $\omega_{it}$ is a series of random shocks when observing capacity.

Defined as daily communication at work, in everyday life, or in other circumstances, in which communicators behave naturally and artlessly (Bell, 2016). In addition, natural communication does not lead to illegal outcomes in the management of village affairs, such as conspiracy or rent-seeking (Baker and Faulkner, 1993). As communication happens often and in various situations, communicators can hardly, and are thus unwilling to, behave strategically to hide their true personal characteristics. Therefore, it is acceptable to assume that first, the times that natural communication occurs is sufficiently large, that is, $T^{Ele (App)} \to +\infty$, and

---

10 The natural communication of the $t$-th time is equivalent to the $t$-th occurrence of natural communication.

11 Times is equivalent to numbers of occurrences.
second, that the series of observations on virtue or capacity is normally distributed.

ASSUMPTION 1 (Natural communication and observations on virtue and capacity): *The representative village resident or township official communicates with the village leader candidates naturally, thus we have* \( v_{it} \sim N(0, \sigma_{\nu}^2) \) and \( \omega_{it} \sim N(0, \sigma_{\omega}^2) \), *and often, thus* \( T^{E \ell e} \) and \( T^{A p p} \) *are sufficiently large.*

Based on Assumption 1, we have \( \Omega_{it}^2 \sim N(\alpha_i, \sigma_{\nu}^2) \), and \( \Omega_{it}^\theta \sim N(\theta_i, \sigma_{\omega}^2) \).

Prior to natural communication, the representative village resident and the township official simply have their own prior perceptions of the virtue and capacity of each village leader candidate, whose distribution is discussed below.

ASSUMPTION 2 (Prior distribution of the virtue and capacity of village leader candidates): *The representative village resident or township official’s prior perceptions of the virtue of village leader candidate* \( i \) *are distributed as* \( N(\alpha_i^e, \sigma_{\nu}^{2e}) \), *truncated at* \([0,1]\), *where* \( \alpha_i^e \in [0,1] \). *Their prior perceptions of the capacity of village leader candidate* \( i \) *are distributed as* \( N(\theta_i^e, \sigma_{\omega}^{2e}) \), *truncated at* \([0,1]\), *where* \( \theta_i^e \in [0,1] \). \( \alpha_i^e \) and \( \theta_i^e \), *the prior means, and* \( \sigma_{\nu}^{2e} \) and \( \sigma_{\omega}^{2e}, \) *the prior variances, are known to the representative village resident or township official.*

As virtue and capacity are assumed to be bounded at \([0,1]\), and the representative village resident or township official engages in long-term natural communication with each village leader candidate, their prior perceptions of the virtue or capacity of each village leader candidate are truncated at \([0,1]\).

Based on their prior perceptions and observations in natural communication, the representative village resident or township official obtains inferred perceptions of the virtue of the village leader candidates. According to Bayes’
rule, these posterior perceptions are obtained by iteration, such that the inferred perceptions of virtue at time $t$ depend on the inferred perceptions of virtue at time $t - 1$ and the observations of virtue at time $t$.

Now we introduce the derivation of the density kernel of the posterior distribution of the virtue of village leader candidate $i$. The representative village resident and the township official have identical prior perceptions of the virtue of village leader candidate $i$, whose density kernel is $\gamma(\alpha_i)$. After naturally communicating with village leader candidate $i$ for the first time, the representative village resident or township official updates the density kernel of the posterior distribution of her virtue as

$$p(\alpha_i | \Omega^0_{i1}) = \gamma(\alpha_i) \cdot L(\alpha_i; \Omega^0_{i1}) \tag{3}$$

This posterior distribution at time $t = 1$ is also the previous distribution at time $t = 2$. After natural communication at time $t = 2$, the updated density kernel of the posterior distribution is given by

$$p(\alpha_i | \Omega^0_{i1}, \Omega^0_{i2}) = [\gamma(\alpha_i) \cdot L(\alpha_i; \Omega^0_{i1})] \cdot L(\alpha_i; \Omega^0_{i2})$$

$$= \gamma(\alpha_i) [L(\alpha_i; \Omega^0_{i1}) \cdot L(\alpha_i; \Omega^0_{i2})] \tag{4}$$

Repeating this iteration, after $T$ times of natural communication between the representative village resident or township official and village leader candidate $i$, the density kernel of the posterior distribution of the virtue of village leader candidate $i$ becomes (See Appendix A.1)

$$p(\alpha_i | \Omega^0_{i1} \cdots \Omega^0_{iT}) = \gamma(\alpha_i) \cdot [L(\alpha_i; \Omega^0_{i1}) \cdots L(\alpha_i; \Omega^0_{iT})]$$

$$\propto \exp\left\{-\frac{1}{2} \left[ \frac{1}{S(\alpha_i)} (\alpha_i - S(\alpha_i))^2 \right]\right\} \tag{5}$$

where the posterior mean of the virtue of village leader candidate $i$ is

$$S(\alpha_i) = \frac{\sigma^2_\alpha \alpha_i \alpha_i}{\sigma^2_\alpha \alpha_i + T \sigma^2_\epsilon \alpha_i} + \frac{T \sigma^2_\epsilon \alpha_i}{\sigma^2_\epsilon \alpha_i + T \sigma^2_\epsilon \alpha_i} \tag{6}$$

and the posterior variance of the virtue of village leader candidate $i$ is

$$\Sigma(\alpha_i) = \frac{\sigma^2_\epsilon \alpha_i \sigma^2_\epsilon \alpha_i}{\sigma^2_\epsilon \alpha_i + T \sigma^2_\epsilon \alpha_i} \tag{7}$$
The posterior mean and the posterior variance are dynamic across the times of natural communication. That is, at time $t = 1, \ldots, T$, the posterior means of the virtue are

$$\frac{\sigma_{\theta i}^2}{\sigma_{\omega+\alpha e}^2} \alpha_i^e + \frac{\sigma_{\alpha e}^2}{\sigma_{\omega+\alpha e}^2} \alpha_i, \ldots, \frac{\sigma_{\theta i}^2}{\sigma_{\omega+\alpha e}^2} \alpha_i^e + \frac{T \sigma_{\theta e}^2}{\sigma_{\omega+\alpha e}^2} \alpha_i,$$

and the posterior variances of the virtue are

$$\frac{\sigma_{\theta i}^2 \sigma_{\alpha e}^2}{\sigma_{\omega+\alpha e}^2}, \ldots, \frac{\sigma_{\alpha e}^2 \sigma_{\theta i}^2}{\sigma_{\omega+\alpha e}^2}.$$

Given the different ultimate times of natural communication in election and in appointment, the representative village resident (by election) or the representative township official (by appointment) obtains the posterior mean of the virtue of village leader candidate $i$ as

$$S_{Ele}^E (App) (\alpha_i) = \frac{\sigma_{\theta i}^2}{\sigma_{\omega+\alpha e}^2} \alpha_i^e + \frac{T \sigma_{\theta e}^2}{\sigma_{\omega+\alpha e}^2} \alpha_i$$

and obtains the posterior variance of the virtue of village leader candidate $i$ as

$$\Sigma_{Ele}^E (App) (\alpha_i) = \frac{\sigma_{\theta i}^2 \sigma_{\alpha e}^2}{\sigma_{\omega+\alpha e}^2}.$$

Following a similar Bayesian inference, the density kernel of the posterior distribution of the capacity of village leader candidate $i$ is given by (See Appendix A.2)

$$p(\theta_i | \Omega_{11}^\theta, \ldots, \Omega_{iT}^\theta) = \gamma(\theta_i) \cdot [L(\theta_i; \Omega_{11}^\theta) \ldots L(\theta_i; \Omega_{iT}^\theta)]$$

$$\propto \exp \left\{ -\frac{1}{2} \sum \left( \theta_i - S(\theta_i) \right)^2 \right\}$$

where the posterior mean of the capacity of village leader candidate $i$ is

$$S(\theta_i) = \frac{\sigma_{\omega}^2 \theta_{\omega}^e + \sigma_{\alpha e}^2 \theta_{\alpha e}^e}{\sigma_{\omega+\alpha e}^2} \theta_i + \frac{T \sigma_{\omega e}^2}{\sigma_{\omega+\alpha e}^2} \theta_i$$

and the posterior variance of the capacity of village leader candidate $i$ is

$$\Sigma(\theta_i) = \frac{\sigma_{\omega}^2 \sigma_{\alpha e}^2 \theta_{\omega}^e + \sigma_{\alpha e}^2 \sigma_{\omega}^2 \theta_{\alpha e}^e}{\sigma_{\omega+\alpha e}^2 \sigma_{\omega+\alpha e}^2}.$$

The posterior mean and the posterior variance are dynamic across the times of natural communication. That is, at time $t = 1, \ldots, T$, the posterior means of
the capacity are \( \left\{ \frac{\sigma^2_{\omega \theta}}{\sigma^2_{\omega \theta} + \sigma^2_{\theta e}} \theta^e_i + \frac{\sigma^2_{\theta e}}{\sigma^2_{\omega \theta} + \sigma^2_{\theta e}} \theta_i, \ldots, \frac{\sigma^2_{\omega \theta}}{\sigma^2_{\omega \theta} + T \sigma^2_{\theta e}} \theta^e_i + \frac{T \sigma^2_{\theta e}}{\sigma^2_{\omega \theta} + T \sigma^2_{\theta e}} \theta_i \right\} \), and the posterior variances of the capacity are \( \left\{ \frac{\sigma^2_{\theta e}}{\sigma^2_{\omega \theta} + \sigma^2_{\theta e}}, \ldots, \frac{\sigma^2_{\theta e}}{\sigma^2_{\omega \theta} + T \sigma^2_{\theta e}} \right\} \).

Given the different ultimate times of natural communication in election and in appointment, the representative village resident (by election) or the representative township official (by appointment) obtains the posterior mean of the capacity of village leader candidate \( i \) as

\[
S^{\text{Ele}} (\text{App}) (\theta_i) = \frac{\sigma^2_{\omega \theta}}{\sigma^2_{\omega \theta} + T^{\text{Ele}} (\text{App}) \sigma^2_{\theta e}} \theta^e_i + \frac{T^{\text{Ele}} (\text{App}) \sigma^2_{\theta e}}{\sigma^2_{\omega \theta} + T^{\text{Ele}} (\text{App}) \sigma^2_{\theta e}} \theta_i
\]

and their posterior variance of the capacity of village leader candidate \( i \) is

\[
\Sigma^{\text{Ele}} (\text{App}) (\theta_i) = \frac{\sigma^2_{\theta e} \sigma^2_{\omega \theta}}{\sigma^2_{\omega \theta} + T^{\text{Ele}} (\text{App}) \sigma^2_{\theta e}}
\]

Proposition 1 explains how the accumulation of natural communication improves the inferences of the virtue and capacity of village leader candidates.

PROPOSITION 1: As the times that the representative village resident or township official communicates naturally with the village leader candidates accumulates, their inference of each village leader candidate’s virtue and capacity is improved in the following aspects:

(a) Inference Precision increases with the times of natural communication, as evidenced by the decrease in the posterior variance of virtue (or capacity) with the ultimate times of natural communication.

(b) Inference Accuracy increases with the times of natural communication, as evidenced by the decrease in the scale ratio of the difference between the posterior mean of virtue (or capacity) and the real value of virtue (or capacity) over the difference between the prior mean of virtue (or capacity) and the real value of virtue (or capacity) with the ultimate times of natural communication.
(c) Marginal Inference Accuracy decreases with the times of natural communication, as evidenced by the increase in the marginal scale ratio of the difference between the posterior mean of virtue (or capacity) and the real value of virtue (or capacity) over the difference between the prior mean of virtue (or capacity) and the real value of virtue (or capacity) with the ultimate times of natural communication.

Proof: (a) The first order derivative of $\Sigma(\alpha_i)$ with respect to $t$ is

$$\frac{\partial[\Sigma(\alpha_i)]}{\partial t} = \frac{-\sigma_{\alpha}^4 e^{\alpha_i^2}}{(\sigma_{\omega+\alpha}^2 + t \sigma_{\omega}^2)^2} < 0$$

Similarly, the first order derivative of $\Sigma(\theta_i)$ with respect to $t$ is

$$\frac{\partial[\Sigma(\theta_i)]}{\partial t} = \frac{-\sigma_{\omega}^4 e^{\theta_i^2}}{(\sigma_{\omega+\omega}^2 + t \sigma_{\omega}^2)^2} < 0$$

(b) The scale ratio of the difference between the posterior mean of virtue and the real value of virtue over the difference between the prior mean of virtue and the real value of virtue is

$$S(\alpha_i) = \frac{\sigma_{\alpha}^2}{(\alpha_i^2 - \alpha_i)} = \frac{\Sigma(\alpha_i)}{\sigma_{\alpha}^2}$$

Therefore, we obtain

$$\frac{\partial[S(\alpha_i)]}{\partial t} = \frac{1}{\sigma_{\alpha}^2} \frac{\partial[\Sigma(\alpha_i)]}{\partial t} < 0$$

Similarly, the scale ratio of the difference between the posterior mean of capacity and the real value of capacity over the difference between the prior mean of capacity and the real value of capacity is

$$S(\theta_i) = \frac{\sigma_{\omega}^2}{(\theta_i^2 - \theta_i)} = \frac{\Sigma(\theta_i)}{\sigma_{\omega}^2}$$

Therefore, we obtain
(20) \[
\frac{\partial}{\partial t} \left[ \frac{S(\theta_i) - \theta_i}{(\theta_i - \theta_i) \sigma_{\theta_i}} \right] = \frac{1}{\sigma_{\theta_i}^2} \frac{\partial S(\theta_i)}{\partial t} < 0
\]

(c) The second order derivative of \( \Sigma(\alpha_i) \) with respect to \( t \) is

(21) \[
\frac{\partial^2 \Sigma(\alpha_i)}{\partial t^2} = \frac{2 \sigma^2_{\alpha_i} \sigma_{\alpha_i}^2}{(\sigma_{\alpha_i}^2 + t \sigma_{\theta_i}^2)^2} > 0
\]

Therefore, the second order derivative of \( \frac{S(\alpha_i) - \alpha_i}{(\alpha_i - \alpha_i)} \) with respect to \( t \), which is the marginal scale ratio of the difference between the posterior mean of virtue and the real value of virtue over the difference between the prior mean of virtue and the real value of virtue, is

(22) \[
\frac{\partial^2 \Sigma(\alpha_i)}{\partial t^2} = \frac{1}{\sigma_{\alpha_i}^2} \frac{\partial^2 S(\alpha_i)}{\partial t^2} > 0
\]

Similarly, the second order derivative of \( \Sigma(\theta_i) \) with respect to \( t \) is

(23) \[
\frac{\partial^2 \Sigma(\theta_i)}{\partial t^2} = \frac{2 \sigma^2_{\theta_i} \sigma_{\theta_i}^2}{(\sigma_{\theta_i}^2 + t \sigma_{\theta_i}^2)^2} > 0
\]

Therefore, the second order derivative of \( \frac{S(\theta_i) - \theta_i}{(\theta_i - \theta_i)} \) with respect to \( t \), which is the marginal scale ratio of the difference between the posterior mean of capacity and the real value of capacity over the difference between the prior mean of capacity and the real value of capacity, is

(24) \[
\frac{\partial^2 \Sigma(\theta_i)}{\partial t^2} = \frac{1}{\sigma_{\theta_i}^2} \frac{\partial^2 S(\theta_i)}{\partial t^2} > 0
\]

[Figure 1 is about here]

Figure 1 shows that as the times that the representative village resident or township official communicates naturally with the village leader candidates accumulates, (a) the bandwidth of the square root of the posterior variance
decreases, reflecting higher inference precision; (b) the difference between the posterior mean of virtue (or capacity) and the real value of virtue (or capacity) decreases, reflecting higher inference accuracy; and (c) the curve of the posterior mean is concave ascending and convex descending, reflecting reduced marginal inference accuracy.

The implication of the improved precision and accuracy of inferences is that each natural communication brings local information on the virtue and capacity of the village leader candidates, leading to more precise and accurate inferences of their virtue and capacity. For marginal inference accuracy, the implication is that as the times of natural communication accumulates, the increment of local information to infer the virtue and capacity of the village leader candidates decreases.

3. Institutional Comparison (Inference Accuracy and Precision): To compare inference precision and inference accuracy after and before the introduction of local direct elections, we make an assumption about $T^{Ele}$ and $T^{App}$, the ultimate times of natural communication between the representative village resident and the representative township official with each village leader candidate.

ASSUMPTION 3: $T^{Ele} > T^{App}$

Assumption 3 indicates that the representative village resident naturally communicates with the village leader candidates more often than the representative township official, implying that the representative village resident has an advantage in terms of local information about the village leader candidates. The reason is that the village leader candidates are also residents of the representative village, with long-term and frequent natural communication with other village residents in various situations (Bell, 2016). For instance, the village leader candidates and other village residents usually know each other
since childhood. As they grow up and live in the same village, they communicate frequently at school, in production or commercial activities, and in everyday life. In contrast, they have fewer opportunities to communicate naturally with township officials. The reasons may be that the village leader candidates usually communicate with township officials when dealing with the public affairs of the village or their private affairs related to township administration, and that township officials are often posted across different towns.

Given that $T_{Ele} > T_{App}$, we compare the precision and accuracy of the inferences of the virtue and capacity of each village leader candidate, after and before the introduction of local direct elections:

(a) The representative village resident’s inferences regarding the virtue and capacity of the village leader candidates are more precise before electing one of them as village leader, compared with those of the representative township official before appointing one of them as village leader.

(b) The representative village resident’s inferences regarding the virtue and capacity of the village leader candidates are more accurate before electing one of them as village leader, compared with those of the representative township official before appointing one of them as village leader.

As shown in Figure 1, the bandwidth representing the square root of the posterior variance of the virtue or capacity of the village leader candidates by election is lower than that by appointment, implying higher inference precision from the representative village resident. The difference between the posterior mean of the virtue or capacity of the village leader candidates and its real value by election is smaller than that by appointment, implying higher inference accuracy from the representative village resident.

In summary, the representative village resident, by naturally communicating more often with the village leader candidates, has more local information about their virtue and capacity than the representative township official. Therefore, as
local direct elections allow the representative village resident and not the representative township official to select the village leader, the virtue and capacity of the village leader candidates are inferred with higher precision and accuracy.

This theoretical demonstration applies to all village committee members. The inferences for each village committee member, including the village leader as representative, are homogenous. Accordingly, the virtue and capacity of each village committee member are inferred more accurately and precisely.

**B. Selection of Village Committee Members**

In this section, we discuss how local direct elections facilitate the meritocratic selection of village committee members through improved *effectiveness of inference*. We discuss the village leader as the representative of all village committee members, our theoretical findings applying to other village committee members. We find that local direct elections in a representative village, by providing more accurate inferences about the virtue and capacity of village leader candidates, improve the expectation of the composite virtue-and-capacity of the village leader. In addition, providing more precise inferences about the virtue and capacity of village leader candidates reduces the variance of the composite virtue-and-capacity of the village leader.

1. **Setup:** The representative village resident and township official both select the village leader candidate with the highest composite virtue-and-capacity as the village leader. Our theory defines composite virtue-and-capacity as a weighted average of the virtue and capacity of village leader candidate $i$, such that $\mu \alpha_i + (1 - \mu) \theta_i$.

$\mu$ represents the weight assigned by the representative village resident and township official to the virtue of the village leader candidates, and $\mu \in [0,1]$. The implication of $\mu$ is the village leader’s virtue-capacity spectrum, which is contingent on the village leader’s tasks and responsibility. It is stipulated that
the village leader be not only capable of developing the local economy, but also abide the laws and be intrinsically motivated to serve the people (National People’s Congress of China, 1998). Therefore, we may assume the village leader’s $\mu$ to be, for example, 0.6. In public sectors, we may assume the politicians’ virtue-capacity spectrum to be $\mu \in (0.5, 1]$. In private sectors, on the contrary, we may assume the leaders’ virtue-capacity spectrum to be $\mu \in [0, 0.5)$, because private sectors have less public purposes but tend to emphasize making profits.

**Expectation of Composite Virtue-and-capacity.** To calculate $[\mu \alpha + (1 - \mu) \theta]^\text{VL.Ele (App)}$, the expectation of the composite virtue-and-capacity of the elected (or appointed) village leader in a representative village, we calculate $\mathbb{E}^\text{Ele (App)}(\mu \alpha_i + (1 - \mu) \theta_i)$, the weighted average of the composite virtue-and-capacity of all village leader candidates in a representative village that has already (or has not) introduced local direct elections, with weight $A_i$ as the probability that village leader candidate $i$ will be elected (or appointed). $A_i$ has the following properties:

(a) $A_i$ is contingent on the composite virtue-and-capacity of village leader candidate $i$.

(b) $A_i \in [0, 1]$, thus its value represents the probability of electing or appointing village leader candidate $i$ as village leader.

(c) $A_i$ is positively associated with the virtue and capacity of village leader candidate $i$, which reflects a positive screening of the election and appointment of village leaders (Dal Bó et al., 2017) based on the candidates’ virtue and capacity. Specifically, $\frac{\partial A_i}{\partial \alpha_i} > 0$ and $\frac{\partial A_i}{\partial \theta_i} > 0$.

To satisfy these three properties, for simplicity and without losing generality, we assume that village leader candidate $i$’s probability of being elected or appointed as village leader increases linearly with a weighted average of her posterior mean of virtue and her posterior mean of capacity, with weight $\mu$, the village leader’s virtue-capacity spectrum. Following Alesina and Tabellini
(2007), the probability of being elected or appointed can be considered a reward. Therefore, the probability that village leader candidate \( i \) will be elected or appointed is given by

\[
A_i = l[\mu S(\alpha_i) + (1 - \mu)S(\theta_i)] \\
= \mu l[\frac{\Sigma(\alpha_i)}{\sigma^2 \alpha} \alpha_i^e + (1 - \frac{\Sigma(\alpha_i)}{\sigma^2 \alpha})\alpha_i] + (1 - \mu)l[\frac{\Sigma(\theta_i)}{\sigma^2 \theta} \theta_i^e + (1 - \frac{\Sigma(\theta_i)}{\sigma^2 \theta})\theta_i]
\]

where \( \Sigma(\alpha_i) \equiv \frac{\sigma^2 \alpha \sigma^2 \alpha}{\sigma^2 \alpha + \tau \sigma^2 \alpha} \), \( \Sigma(\theta_i) \equiv \frac{\sigma^2 \theta \sigma^2 \theta}{\sigma^2 \theta + \tau \sigma^2 \theta} \), and \( l \in [0, 1] \).

As discussed in Section III.A., \( \Sigma(\alpha_i) \) and \( \Sigma(\theta_i) \) measure inference precision. As the times of natural communication increases, \( \Sigma(\alpha_i) \) and \( \Sigma(\theta_i) \) decrease, thus \( A_i \) tends to be the composite virtue-and-capacity. Therefore, when \( T = T^\text{Ele} \), \( \Sigma(\alpha_i) = \Sigma^\text{Ele}(\alpha_i) \), and \( \Sigma(\theta_i) = \Sigma^\text{Ele}(\theta_i) \), we have \( A_i = A_i^\text{Ele} \), the probability that village leader candidate \( i \) will be elected. When \( T = T^\text{App} \), \( \Sigma(\alpha_i) = \Sigma^\text{App}(\alpha_i) \), and \( \Sigma(\theta_i) = \Sigma^\text{App}(\theta_i) \), we have \( A_i = A_i^\text{App} \), the probability that village leader candidate \( i \) will be appointed.

When calculating the weighted average of the composite virtue-and-capacity of all village leader candidates with weight \( A_i \), as \( \int_0^1 \int_0^1 A_i d\alpha_i d\theta_i < 1 \), that is, the sum of all weights is less than 1, we should have \( \int_0^1 \int_0^1 [\mu \alpha_i + (1 - \mu)\theta_i] A_i d\alpha_i d\theta_i \), the weighted average of the composite virtue-and-capacity of all village leader candidates, divided by \( \int_0^1 \int_0^1 A_i d\alpha_i d\theta_i \), to standardize the weights. As a result, the expectation of the composite virtue-and-capacity of the elected (or appointed) village leader is given by

\[
[\mu \alpha + (1 - \mu)\theta]^\text{VL, Ele} (\text{App}) = \mathbb{E}^\text{Ele} (\text{App}) (\mu \alpha_i + (1 - \mu)\theta_i) \\
= \frac{\int_0^1 \int_0^1 [\mu \alpha_i + (1 - \mu)\theta_i] A_i^\text{Ele} (\text{App}) d\alpha_i d\theta_i}{\int_0^1 \int_0^1 A_i^\text{Ele} (\text{App}) d\alpha_i d\theta_i}
\]

where the probability that village leader candidate \( i \) will be elected (or appointed) is

\[
A_i^\text{Ele} (\text{App}) = \mu S^\text{Ele} (\text{App}) (\alpha_i) + (1 - \mu)S^\text{Ele} (\text{App}) (\theta_i)
\]

\[20\]
\[
\begin{align*}
&= \mu_l \left[ \frac{\Sigma^{E_l}(\theta_l)}{\sigma^{2}_{e}} \alpha^e_l + \left(1 - \frac{\Sigma^{E_l}(\theta_l)}{\sigma^{2}_{e}}\right) \alpha_l \right] \\
&\quad+ (1 - \mu_l) \left[ \frac{\Sigma^{E_l}(\theta_l)}{\sigma^{2}_{e}} \theta^e_l + \left(1 - \frac{\Sigma^{E_l}(\theta_l)}{\sigma^{2}_{e}}\right) \theta_l \right]
\end{align*}
\]

where \( \Sigma^{E_l}(\alpha_l) \equiv \frac{\sigma^{2}_{\alpha} \sigma^{2}_{\alpha} + \sigma^{2}_{\alpha} \sigma^{2}_{\alpha}}{\sigma^{2}_{\alpha} + \sigma^{2}_{\alpha}} \), \( \Sigma^{E_l}(\theta_l) \equiv \frac{\sigma^{2}_{\theta} \sigma^{2}_{\theta} + \sigma^{2}_{\theta} \sigma^{2}_{\theta}}{\sigma^{2}_{\theta} + \sigma^{2}_{\theta}} \), and \( l \in [0,1] \). Therefore, this shows that what distinguishes the expectation of the composite virtue-and-capacity of the elected village leader and the appointed village leader in a representative village is the times that the representative village resident and township official communicate naturally with the village leader candidates.

**Variance of the Composite Virtue-and-capacity.** We can also calculate \( \text{Var}^{VL,E_l}(\mu \alpha + (1 - \mu) \theta) \), the variance of the composite virtue-and-capacity of the elected (or appointed) village leader in a representative village. To this end, we calculate \( \text{Var}^{E_l}(\mu \alpha_i + (1 - \mu) \theta_i) \), the variance of the composite virtue-and-capacity of all village leader candidates in a representative village that has already (or has not) introduced local direct elections. By definition, the variance of the composite virtue-and-capacity of the elected (or appointed) village leader in a representative village is

\[
\text{Var}^{VL,E_l}(\mu \alpha + (1 - \mu) \theta) = \text{Var}^{E_l}(\mu \alpha_i + (1 - \mu) \theta_i)
\]

\[
= \mathbb{E}^{E_l}(\mu \alpha_i + (1 - \mu) \theta_i)^2 - \mathbb{E}^{E_l}(\mu \alpha_i + (1 - \mu) \theta_i)^2
\]

This measures to what extent the composite virtue-and-capacity of the elected (or appointed) village leader varies. Similar to the expectation, this shows that what distinguishes the variance of the composite virtue-and-capacity of the elected village leader and the appointed village leader in a representative village is the times that the representative village resident and township official communicate naturally with the village leader candidates.

**2. Institutional Comparison (Meritocratic Selection of Village Leaders):**

Proposition 2 compares \( [\mu \alpha + (1 - \mu) \theta]^{VL,E_l} \), the expectation of the composite virtue-and-capacity of the elected village leader, with \( [\mu \alpha + (1 - \mu) \theta]^{VL,App} \), the expectation of the composite virtue-and-capacity of the
appointed village leader.

PROPOSITION 2: The expectation of the composite virtue-and-capacity of the elected village leader is greater than that of the appointed village leader in a representative village. Specifically, we have

\[(\mu \alpha + (1 - \mu)\theta)^{V.L,Ele} > (\mu \alpha + (1 - \mu)\theta)^{V.L,App}\]

where \((\mu \alpha + (1 - \mu)\theta)^{V.L,Ele} \equiv E^{Ele}(\mu \alpha_i + (1 - \mu)\theta_i)\) and \((\mu \alpha + (1 - \mu)\theta)^{V.L,App} \equiv E^{App}(\mu \alpha_i + (1 - \mu)\theta_i)\). Here are some specific cases:

Case 1: \(\mu = 1\); that is, the representative village resident or township official only considers the virtue of the village leader candidates.

Case 2: \(\mu = 0\); that is, the representative village resident or township official only considers the capacity of the village leader candidates.

Case 3: \(\mu = \frac{1}{2}\); that is, the representative village resident or township official considers the virtue and capacity of the village leader candidates with equal weights.

Case 4: \(\mu \in (\frac{1}{2}, 1)\); that is, the representative village resident or township official puts more emphasis on virtue. This is contingent on \(\theta^e_i \in [0.5, 1]\). that is, the prior mean of the capacity of each village leader candidate is greater than the mean of the real value of the capacity of all potential village leader candidates.

Case 5: \(\mu \in (0, \frac{1}{2})\); that is, the representative village resident or township official puts more emphasis on capacity. This is contingent on \(\alpha^e_i \in [0.5, 1]\), that is, the prior mean of the capacity of each village leader candidate is greater than the mean of the real value of the capacity of all potential village leader candidates.

Proof: See Appendix B.1.

In practice, both Case 4 and Case 5 hold, that is, the conditions in both cases exist all the time. According to the requirements of the OLVC, village leader candidates satisfy certain personal characteristics in terms of capacity, such as
a diploma or management experience (National People’s Congress of China, 1998). In this sense, the mean of the prior perceptions of the representative village resident or township official regarding the capacity of each village leader candidate is greater than or equal to 0.5, the mean of the capacity of all village residents, in other words, $\theta_i^p \in [0.5, 1]$. The OLVC also requires that village leader candidates satisfy certain personal characteristics in terms of virtue, such as no criminal record or affiliation to the CPC (National People’s Congress of China, 1998). In this sense, the mean of the prior perceptions of the representative village resident or township official regarding the virtue of the village leader candidates is greater than or equal to 0.5, the mean of the virtue of all village residents.

Proposition 3 compares $\text{Var}^{\text{VL, Ete}}(\mu\alpha + (1 - \mu)\theta)$, the variance of the composite virtue-and-capacity of the elected village leader, with $\text{Var}^{\text{VL, App}}(\mu\alpha + (1 - \mu)\theta)$, the variance of the composite virtue-and-capacity of the appointed village leader.

**PROPOSITION 3:** The variance of the composite virtue-and-capacity of the elected village leader is smaller than that of the appointed village leader in a representative village. Specifically, we have

$$(30) \quad \text{Var}^{\text{VL, Ete}}(\mu\alpha + (1 - \mu)\theta) < \text{Var}^{\text{VL, App}}(\mu\alpha + (1 - \mu)\theta)$$

whenever the value of $\mu$ in $[0, 1]$.

It should be noted, as discussed in (28), that $\text{Var}^{\text{VL, Ete}}(\mu\alpha + (1 - \mu)\theta) = \text{Var}^{\text{Ete}}(\mu\alpha_i + (1 - \mu)\theta_i)$, and $\text{Var}^{\text{VL, App}}(\mu\alpha + (1 - \mu)\theta) = \text{Var}^{\text{App}}(\mu\alpha_i + (1 - \mu)\theta_i)$.

Proof: See Appendix B.2. $\blacksquare$

In summary, after the introduction of local direct elections in a representative village, the expectation of the composite virtue-and-capacity of the village leader increases while its variance decreases. This implies that by providing
local information on the virtue and capacity of village leader candidates, local direct elections facilitate the meritocratic selection of village leaders.

This theoretical demonstration applies to all village committee members. The selection of each village committee member, including the village leader as representative, is homogenous. Thus, the expectation of the composite virtue-and-capacity of each village committee member is also homogenous. Accordingly, by providing local information on the virtue and capacity of village committee candidates, local direct elections increase the expectation of the composite virtue-and-capacity of each village committee member, while reducing its variance. In other words, local direct elections facilitate the meritocratic selection of all village committee members.

C. Selection of Village Party Secretaries

In this section, we discuss how local direct elections facilitate the meritocratic selection of village party secretaries through improved candidate pools. The inferences of the village party secretary candidates and the selection of the village party secretary are similar to those discussed in Sections III.A. and III.B. Part of or all village committee members, including village leaders, are also village party committee members and are therefore candidates for village party secretaries. Therefore, local direct elections in a representative village, by increasing the expectation of the composite virtue-and-capacity of village committee members, indirectly improve the expectation of the composite virtue-and-capacity of the village party secretary.

1. Performance-based Promotion of Village Party Secretaries: In a representative village, village party committee members are also the candidates for the village party secretary. The representative township official, by observing the performance of each village party committee member, infers their composite virtue-and-capacity and appoints the candidate with the highest composite as village party secretary, namely the performance-based promotion of the village party secretary.
Village party committee members work on a series of tasks on village affairs. Contingent on these tasks, each village party committee member has her virtue-capacity spectrum. We have the following assumption for village party committee members’ virtue-and-capacity spectrum:

ASSUMPTION 4 (Village party committee members’ virtue-capacity spectrum): Village party committee members’ virtue-capacity spectrum are also μ, same as village committee members’ virtue-capacity spectrum.

The implication of Assumption 4 is that because both village committee members and village party committee members work in the public sector, and they both work on village affairs, it is assumed that village party committee members’ tasks require the same ratio of virtue over capacity as village committee members tasks.

Therefore, similar to Section III.A., it is assumed that the performance of village party committee member \( j \) in task \( m = 1, 2, ..., M \) is the sum of her composite virtue-and-capacity and a random error, following Jones and Olken (2005), Besley et al. (2011), Yao and Zhang (2015), and Bloom et al. (2015). Specifically,

\[
P_{jm} = \mu \alpha_j + (1 - \mu) \theta_j + \epsilon_{jm}
\]

where \( P_{jm} \) represents the performance of village party committee member \( j \) in task \( m = 1, 2, ..., M \), \( M \) being sufficiently large. \( \mu \alpha_j + (1 - \mu) \theta_j \) represents her composite virtue-and-capacity with her virtue-and-capacity spectrum \( \mu \). \( \epsilon_{jm} \) represents a series of random shocks, where \( \epsilon_{jm} \sim N(0, \sigma^2_{\epsilon}) \). As a result, \( P_{jm} \sim N(\mu \alpha_j + (1 - \mu) \theta_j, \sigma^2_{\epsilon}) \), similar to Section III.A.

Similar to Section III.B., the probability of appointing village party committee member \( j \) as village party secretary increases linearly with the posterior mean of her composite virtue-and-capacity. Specifically,

\[
R_j = wS[\mu \alpha_j + (1 - \mu) \theta_j]
\]
where $S[\mu \alpha_j + (1 - \mu) \theta_j]$ is the posterior mean of the composite virtue-and-capacity of village party committee member $j$ by observing her performance in a series of tasks $m = 1, 2, ..., M$. $\omega \in [0,1]$.

To specify $S[\mu \alpha_j + (1 - \mu) \theta_j]$, we must first assume the prior perception distribution of the representative township official of the composite virtue-and-capacity of an arbitrary village party committee member.

ASSUMPTION 5 (Prior distribution of the virtue and capacity of village party committee members): The prior perception of the representative township official regarding the virtue of an arbitrary village party committee member is distributed as $N(\alpha^u, \sigma_{\alpha^u}^2)$, truncated at $[0,1]$, where $\alpha^u \in [0,1]$. Her prior perception of the capacity of an arbitrary village party committee member is distributed as $N(\theta^u, \sigma_{\theta^u}^2)$, truncated at $[0,1]$, where $\theta^u \in [0,1]$. $\alpha^u$ and $\theta^u$, the prior means, and $\sigma_{\alpha^u}^2$ and $\sigma_{\theta^u}^2$, the prior variances, are known to the representative township official.

The prior distribution of the composite virtue-and-capacity of an arbitrary village party committee member is thus a normal distribution $N(\mu \alpha^u + (1 - \mu) \theta^u, \mu^2 \sigma_{\alpha^u}^2 + (1 - \mu)^2 \sigma_{\theta^u}^2)$, truncated at $[0,1]$. Using the same Bayesian inference framework as in Section III.A., after $M$ iterations, the posterior mean of the composite virtue-and-capacity of village party committee member $j$ is given by (See Appendix C)

\begin{equation}
S[\mu \alpha_j + (1 - \mu) \theta_j] = (1 - \varphi)[\mu \alpha^u + (1 - \mu) \theta^u] + \varphi[\mu \alpha_j + (1 - \mu) \theta_j]
\end{equation}

where $\mu \in (0,1)$ and $\varphi \equiv \frac{M[\mu^2 \sigma_{\alpha^u}^2 + (1 - \mu)^2 \sigma_{\theta^u}^2]}{\sigma^2 + M[\mu^2 \sigma_{\alpha^u}^2 + (1 - \mu)^2 \sigma_{\theta^u}^2]}$.

Therefore, the probability of appointing village party committee member $j$ as village party secretary is

\begin{equation}
R_j = \omega S[\mu \alpha_j + (1 - \mu) \theta_j]
\end{equation}
This represents the performance-based promotion; that is, the probability that village party committee member \( j \) will be promoted to village party secretary is positively associated with her composite virtue-and-capacity, which is inferred by observing her performance in a series of tasks.

2. Institutional Comparison (Meritocratic Selection of Village Party Secretaries): The village party committee has two types of members: (I) individuals who are both members of the village party committee and the village committee, denoted by village party committee member \( j_k, j_k = \{1, 2, \ldots\} \); and (II) individuals who are simply members of the village party committee, denoted by village party committee member \( j_k, j_k = \{1, 2, \ldots\} \). The village leader belongs to Type I village party committee members.

The candidate pool for the village party secretary is partially improved by local direct elections. After the introduction of local direct elections, the expectation of the composite virtue-and-capacity of Type I village party committee members increases, while that of Type II village party committee members remains unchanged. Indeed, with local direct elections, Type I village party committee members are no longer appointed by the representative township official, but elected by the representative village resident, thus the expectation of their composite virtue-and-capacity increases, as discussed in Section III.B. In contrast, as Type II village party committee members remain appointed by the representative township official, their composite virtue-and-capacity remains unchanged. On the contrary, before the introduction of local direct elections, the expectations of the composite virtue-and-capacity of both Type I and Type-II village party committee members were the same, because both Type I and Type-II village party committee members were appointed by the representative township official.

Proposition 4 explains how local direct elections facilitate the meritocratic selection of the village party secretary because of the partially improved candidate pool. Let the expectation of the composite virtue-and-capacity of
Type I village party committee members be \([\mu \alpha_{j_k} + (1 - \mu)\theta_{j_k}]\) and their probability of promotion be \(R_{j_k}\), where \(R_{j_k} = (1 - \varphi)w[\mu\alpha^u + (1 - \mu)\theta^u] + \varphi w[\mu\alpha_{j_k} + (1 - \mu)\theta_{j_k}],\) according to (34). Therefore, the expectation of the composite virtue-and-capacity of elected (or appointed) Type I village party committee members is \([\mu\alpha_{j_k} + (1 - \mu)\theta_{j_k}]^{Ele}^{(App)}\) and their probability of promotion is \(R_{j_k}^{Ele}^{(App)} = (1 - \varphi)w[\mu\alpha^u + (1 - \mu)\theta^u] + \varphi w[\mu\alpha_{j_k} + (1 - \mu)\theta_{j_k}]^{Ele}^{(App)}\). In addition, let the expectation of the composite virtue-and-capacity of Type II village party committee members be \([\mu\alpha_{j_k} + (1 - \mu)\theta_{j_k}]\) and their probability of promotion be \(R_{j_k}^\prime\), where \(R_{j_k}^\prime = (1 - \varphi)w[\mu\alpha^u + (1 - \mu)\theta^u] + \varphi w[\mu\alpha_{j_k} + (1 - \mu)\theta_{j_k}],\) according to (34). As discussed earlier, \([\mu\alpha_{j_k} + (1 - \mu)\theta_{j_k}]\) is not correlated with the introduction of local direct elections. Finally, let \([\mu\alpha + (1 - \mu)\theta]^{VPS}\) be the expectation of the composite virtue-and-capacity of the village party secretary, where \([\mu\alpha + (1 - \mu)\theta]^{VPS,After}\) is the expectation of the composite virtue-and-capacity of the village party secretary directly elected as a village committee member (i.e., starting her political career after the introduction of local direct elections), and \([\mu\alpha + (1 - \mu)\theta]^{VPS,Before}\) is the expectation of the composite virtue-and-capacity of the village party secretary never directly elected as a village committee member (i.e., she started her political career before the introduction of local direct elections).

PROPOSITION 4 (Meritocratic Selection of Village Party Secretaries): After the introduction of local direct elections in a representative village, the expectation of the composite virtue-and-capacity of the village party secretary increases, due to the increased expectation of the composite virtue-and-capacity of Type I village party committee members. Specifically, we have

\[
[\mu\alpha + (1 - \mu)\theta]^{VPS,After} > [\mu\alpha + (1 - \mu)\theta]^{VPS,Before}
\]

That is, the expectation of the composite virtue-and-capacity of the village party secretary directly elected as a village committee member is greater than that of
the village party secretary who has never been elected directly as a village committee member.

Proof: See Appendix D. ■

D. Summary

This section demonstrates that local direct elections, by providing local information, facilitate meritocratic selection both for village committee members and village party secretaries, but in different ways.

Local direct elections directly facilitate the meritocratic selection of village committee members, by improving the effectiveness of inference. After local direct elections transfer the authority of selecting village committee members from township officials to village residents, who have an advantage regarding local information about these candidates, the virtue and capacity of these candidates are inferred with higher precision and accuracy. As a result, the expectation of the composite virtue-and-capacity of village committee members increases while its variance decreases. With the virtue-capacity spectrum in our model, our demonstrations apply to a continuum of selection scenarios, from purely virtue-based to purely capacity-based.

In contrast, local direct elections indirectly facilitate the meritocratic selection of village party secretaries, by improving the candidate pools. As part of or all village committee members, including village leaders, are also village party committee members, they are also candidates for village party secretaries. As local direct elections increase the expectation of the composite virtue-and-capacity of village party committee members, the expectation of the composite virtue-and-capacity of village party secretaries also increases.

IV. Concluding Remarks

This paper, to our knowledge, is the first to use local information to address the adverse selection in political selection. Two major problems of political economy, both related to information asymmetry, are political selection that
suffers from adverse selection and political incentive that suffers from moral hazard. Many studies address moral hazard, either by explaining current institutional arrangements or by designing new mechanisms, to discuss the incentive of politicians (Laffont, 2000; Besley, 2006). However, few studies discuss the adverse selection in the selection of politicians (Besley, 2005). The adverse selection, on the contrary, is commonly discussed, for example, in the literature on job market signaling (Spence, 1973) and product advertising (Milgrom and Roberts, 1986).

This paper’s theory emphasizes the role of local information in addressing the adverse selection in political selection. It is shown that the accumulating local information, quantified with the increasing numbers of natural communication between political candidates and political decision makers, improve the effectiveness of inference of each political candidate’s virtue and capacity. Specifically, as the numbers of natural communication increase, political decision makers infer each political candidate’s virtue and capacity more accurately and precisely. To frame the theory, this paper uses the Bayesian inference framework instead of the game theory framework, because this paper focuses on the influences of accumulating local information on the effectiveness of inference rather than the strategic behaviors of political candidates and decision makers.

The essential mechanism through which local direct elections work in the meritocratic selection of politicians is the local information on the political candidates that local direct elections provide. In the context of Chinese local governance, the introduction of local direct elections, by providing local information on each village committee candidate’s virtue and capacity, enhances the expectation, while reduces the variance, of the composite virtue-and-capacity of each village committee member. Further, the expectation of the composite virtue-and-capacity of the village party secretary is also enhanced indirectly. In this sense, one of the criteria to evaluate an institutional arrangement’s effect in meritocratic selection is whether that institutional
arrangement provide sufficient local information to infer political candidates’ virtue and capacity.

This paper’s theory, although discussed in the context of Chinese local governance, has general implications. The essential scenario that this paper’s theory discusses is characterized as the small groups in the grassroots level of the stratified governance structure, regardless of rural or urban area, and public or private sector. In this essential scenario, the leader of each small group is locally and directly elected, then has the chance to be promoted upward, based on her performance observed. A typical example scenario is an organization with multiple departments, in which the head of each department is directly elected within each department, then those heads are likely to be promoted into the board of the organization based on their performance. In this process, the local information on each head candidate’s virtue and capacity is aggregated, facilitating the meritocratic selection of each head and the board directors.

Several limitations and extensions should be noted. (1) It is assumed in our theory that the political selection in rural China is law-abiding. A future study may discuss the scenarios like bribes in elections and conspiracy in the performance-based promotion. (2) It is neglected in our theory to discussed whether the meritocratic selection facilitated by the local direct election yields an equilibrium in which village committee members, village party secretaries, village residents, and township officials all obtain optimized gains. A future study may discuss the existence of such an equilibrium, and how that equilibrium deviates when relaxing the assumption of law-abiding in political selection. (3) It is assumed in our theory that political candidates are passive in signaling their virtue and capacity. A future research may discuss political candidates’ positive behaviors in signaling their virtue and capacity to village residents or township officials. (4) Village residents or township officials’ perceptions on each political candidate are assumed homogenous. A future study may discuss how various distributions of village residents or township officials’ perceptions influence the political selection.
REFERENCES


http://news.12371.cn/2015/03/11/ARTI1426061212036535.shtml

Chan, Joseph. “Political Meritocracy and Meritorious Rule: A Confucian Perspective.” *The East Asian Challenge for Democracy: Political Meritocracy in Comparative Perspective* edited by Daniel A. Bell, and
Elman, Benjamin A. *Civil Examinations and Meritocracy in Late Imperial China*. Cambridge, MA: Harvard University Press, 2013.


Figure 1: Inference Accuracy and Precision with Natural Communication Times
Online Appendices for Local Direct Elections, Local Information, and Meritocratic Selection

A. Posterior Distributions of the Virtue and the Capacity of Village Leader Candidates

A.1 Virtue: Following Koop et al. (2007) and Lancaster (2004), we derive the density kernel of the posterior distribution of virtue, such that

\[
p(\alpha_i|\Omega_{it}^\alpha) = \gamma(\alpha_i) \cdot [L(\alpha_i; \Omega_{it}^\alpha) \ldots L(\alpha_i; \Omega_{it}^\alpha) \ldots L(\alpha_i; \Omega_{it}^\alpha)]
\]

where the vector \( \Omega_{it}^\alpha \) is \([\Omega_{it1}^\alpha \ldots \Omega_{itT}^\alpha] \).

First, we derive \( \gamma(\alpha_i) \), the density kernel of the prior distribution of virtue. As the prior distribution is a normal distribution truncated at \([0, 1]\), we have

\[
\gamma(\alpha_i) = \gamma(\alpha_i | 0 \leq \alpha_i \leq 1, \alpha_i^e)
\]

\[
= \frac{1}{\pi \sigma_{\alpha}^2} \cdot \frac{\phi \left( \frac{\alpha_i - \alpha_i^e}{\sigma_{\alpha}^2} \right)}{P(\alpha_i | 0 \leq \alpha_i \leq 1, \alpha_i^e)}
\]

\[
= \frac{(2\pi \sigma_{\alpha}^2)^{-1/2} \exp \left( \frac{-(\alpha_i - \alpha_i^e)^2}{2\sigma_{\alpha}^2} \right)}{P(\alpha_i | 0 \leq \alpha_i \leq 1, \alpha_i^e)}
\]

\[
\propto \frac{\exp \left( \frac{-(\alpha_i - \alpha_i^e)^2}{2\sigma_{\alpha}^2} \right)}{P(\alpha_i | 0 \leq \alpha_i \leq 1, \alpha_i^e)}
\]

where \( P(\alpha_i | 0 \leq \alpha_i \leq 1, \alpha_i^e) = \Phi \left( \frac{1 - \alpha_i^e}{\sigma_{\alpha}^2} \right) - \Phi \left( \frac{-\alpha_i^e}{\sigma_{\alpha}^2} \right) \) represents the probability that \( \alpha_i \) is at \([0, 1]\), contingent on \( \alpha_i^e \).

Second, we derive the joint density of \( \Omega_{it}^\alpha \) as a likelihood function given by

\[
L(\alpha_i; \Omega_{it}^\alpha) = \prod_{t=1}^{T} (2\pi \sigma_{\alpha}^2)^{-1/2} \cdot \exp \left\{ -\frac{\left( \Omega_{it}^\alpha - \alpha_i \right)^2}{2\sigma_{\alpha}^2} \right\}
\]

\[
= \left( 2\pi \sigma_{\alpha}^2 \right)^{-T/2} \cdot \exp \left\{ -\sum_{t=1}^{T} \frac{\left( \Omega_{it}^\alpha - \alpha_i \right)^2}{2\sigma_{\alpha}^2} \right\}
\]

\[
\propto \exp \left\{ -\frac{1}{2\sigma_{\alpha}^2} \sum_{t=1}^{T} \left( \Omega_{it}^\alpha - \alpha_i \right)^2 \right\}
\]

\[
\propto \exp \left\{ -\frac{1}{2\sigma_{\alpha}^2} \sum_{t=1}^{T} \left( \Omega_{it}^\alpha - \Omega_t^\alpha + \Omega_t^\alpha - \alpha_i \right)^2 \right\}
\]

\[
\propto \exp \left\{ -\frac{1}{2\sigma_{\alpha}^2} \left[ \sum_{t=1}^{T} (\Omega_{it}^\alpha - \Omega_t^\alpha)^2 + \sum_{t=1}^{T} (\Omega_t^\alpha - \alpha_i)^2 + 2 \sum_{t=1}^{T} (\Omega_{it}^\alpha - \Omega_t^\alpha)(\Omega_{it}^\alpha - \alpha_i) \right] \right\}
\]

where \( \overline{\Omega}_{it}^\alpha = \frac{1}{T} \sum_{t=1}^{T} \Omega_{it}^\alpha \). As \( \sum_{t=1}^{T} (\Omega_{it}^\alpha - \Omega_t^\alpha)^2 \), the second-order moment, is a constant, \( 2 \sum_{t=1}^{T} (\Omega_{it}^\alpha - \Omega_t^\alpha)(\Omega_{it}^\alpha - \alpha_i) = 2(\overline{\Omega}_{it}^\alpha - \alpha_i) \sum_{t=1}^{T} (\Omega_{it}^\alpha - \overline{\Omega}_{it}^\alpha) = 0 \), we have
(A4) \( L(\alpha_i; \Omega_{it}^\alpha) \propto \exp \left\{ -\sum_{t=1}^T \frac{(\bar{\alpha}_t^\alpha - \alpha_i)^2}{2\sigma_{\alpha t}^2} \right\} \propto \exp \left\{ -T \left( \frac{(\bar{\alpha}_t^\alpha - \alpha_i)^2}{2\sigma_{\alpha t}^2} \right) \right\} \)

Third, the density kernel of the posterior distribution of virtue is

\[
L(\alpha_i; \Omega_{it}^\alpha) \gamma(\alpha_i) \propto \frac{1}{2} \exp \left( \frac{1}{\sigma_{\alpha t}^2} \left( \alpha_i - 2\alpha_i \Omega_{it}^\alpha + \alpha_i^2 \right) + \frac{1}{\sigma_{\alpha e}^2} \left( \alpha_i^2 - 2\alpha_i \alpha_i^2 + \alpha_i^2 \right) \right) \cdot \frac{1}{2} \exp \left( \frac{1}{\sigma_{\alpha t}^2} \left( \alpha_i - 2\alpha_i \Omega_{it}^\alpha + \alpha_i^2 \right) \right) \cdot \frac{1}{2} \exp \left( \frac{1}{\sigma_{\alpha e}^2} \left( \alpha_i^2 - 2\alpha_i \alpha_i^2 + \alpha_i^2 \right) \right)
\]

Completing the square in the numerator of (A5), we have,

(6) \( L(\alpha_i; \Omega_{it}^\alpha) \cdot \gamma(\alpha_i) \propto \frac{1}{2} \exp \left( \frac{1}{\sigma_{\alpha t}^2 + T\sigma_{\alpha e}^2} \left( \alpha_i - \frac{\sigma_{\alpha t}^2 \alpha_t^\alpha}{\sigma_{\alpha t}^2 + T\sigma_{\alpha e}^2} \right)^2 \right) \cdot \frac{1}{2} \exp \left( \frac{1}{\sigma_{\alpha e}^2 + T\sigma_{\alpha e}^2} \left( \alpha_i - \frac{\sigma_{\alpha e}^2 \alpha_e^\alpha}{\sigma_{\alpha t}^2 + T\sigma_{\alpha e}^2} \right)^2 \right) \cdot \frac{1}{2} \exp \left( \frac{1}{\sigma_{\alpha t}^2 + T\sigma_{\alpha e}^2} \left( \alpha_i - \frac{\sigma_{\alpha t}^2 \alpha_t^\alpha}{\sigma_{\alpha t}^2 + T\sigma_{\alpha e}^2} \right)^2 \right)
\]

where \( P(\alpha_i | 0 \leq \alpha_i \leq 1, \alpha_i^e) = \Phi \left( \frac{\alpha_i^e}{\alpha_i^e} \right) - \Phi \left( \frac{-\alpha_i^e}{\alpha_i^e} \right) \). Therefore, this density kernel is still a truncated normal distribution, in which the posterior variance of virtue is

(7) \( \Sigma(\alpha_i) = \frac{\sigma_{\alpha t}^2 \sigma_{\alpha e}^2}{\sigma_{\alpha t}^2 + T\sigma_{\alpha e}^2} \)

and the posterior mean of virtue is

(8) \( S(\alpha_i) = \Sigma(\alpha_i) \left[ \frac{1}{\sigma_{\alpha t}^2} \alpha_t^\alpha + \frac{T}{\sigma_{\alpha t}^2} \Omega_{it}^\alpha \right] = \frac{\sigma_{\alpha t}^2 \sigma_{\alpha e}^2}{\sigma_{\alpha t}^2 + T\sigma_{\alpha e}^2} \alpha_t^\alpha + \frac{T\sigma_{\alpha e}^2}{\sigma_{\alpha t}^2 + T\sigma_{\alpha e}^2} \frac{\bar{\Omega}_i^\alpha}{T} \)

As \( \Omega_{it}^\alpha \sim \mathcal{N}(\alpha_i, \sigma_{\alpha t}^2) \) and \( T = T \text{Ele} (T \text{App}) \rightarrow +\infty \), we have

(9) \( \bar{\Omega}_i^\alpha \equiv \frac{1}{T} \sum_{t=1}^T \Omega_{it}^\alpha = \mathbb{E} (\Omega_{it}^\alpha) = \alpha_i \)

thus, the posterior mean of virtue is

(A10) \( S(\alpha_i) = \frac{\sigma_{\alpha t}^2}{\sigma_{\alpha t}^2 + T\sigma_{\alpha e}^2} \alpha_t^\alpha + \frac{T\sigma_{\alpha e}^2}{\sigma_{\alpha t}^2 + T\sigma_{\alpha e}^2} \alpha_i \)

A.2 Capacity: Similar to virtue, we derive the density kernel of the posterior distribution of capacity, such that
(A11) \[ p(\theta_i|\Omega_{it}^\theta) = \gamma(\theta_i) \cdot \left[ L(\theta_i; \Omega_{it}^\theta) \cdots L(\theta_i; \Omega_{it}^\theta) \right] \]

where the vector \( \Omega_{it}^\theta \) is \( [\Omega_{it}^\theta \cdots \Omega_{it}^\theta \cdots \Omega_{it}^\theta] \).

First, we derive \( \gamma(\theta_i) \), the density kernel of the prior distribution of capacity. As the prior distribution is a normal distribution truncated at \( [0, 1] \), we have

(A12) \[ \gamma(\theta_i) = \gamma(\theta_i|0 \leq \theta_i \leq 1, \theta_i^e) \]

\[ \propto \exp \left( \frac{- (\theta_i - \theta_i^e)^2}{2 \sigma_{\theta i}^2} \right) \]

where \( P(\theta_i|0 \leq \theta_i \leq 1, \theta_i^e) = \Phi \left( \frac{1 - \theta_i^e}{\sigma_{\theta i}^e} \right) - \Phi \left( \frac{-\theta_i^e}{\sigma_{\theta i}^e} \right) \) represents the probability that \( \theta_i \) is at \( [0, 1] \), contingent on \( \theta_i^e \).

Second, we derive the joint density of \( \Omega_{it}^\theta \) as a likelihood function given by

(A13) \[ L(\theta_i; \Omega_{it}^\theta) = \prod_{t=1}^{T} \left( 2\pi \sigma_{\omega \theta}^2 \right)^{-\frac{1}{2}} \cdot \exp \left\{ \frac{- (\theta_i^0 - \theta_i)^2}{2 \sigma_{\omega \theta}^2} \right\} \]

\[ \propto \exp \left\{ - \sum_{t=1}^{T} \left( \frac{\theta_i^0 - \theta_i}{2 \sigma_{\omega \theta}^2} \right)^2 \right\} \]

where \( \Omega_{it}^\theta = \frac{1}{T} \sum_{t=1}^{T} \Omega_{it}^\theta \). As \( \sum_{t=1}^{T} \left( \Omega_{it}^\theta - \Omega_{i}^\theta \right)^2 \), the second-order moment, is a constant,

\[ 2 \sum_{t=1}^{T} (\Omega_{it}^\theta - \Omega_{i}^\theta)(\Omega_{i}^\theta - \theta_i) = 2(\Omega_{i}^\theta - \theta_i) \sum_{t=1}^{T} (\Omega_{it}^\theta - \Omega_{i}^\theta) = 0, \]

we have

(A14) \[ L(\theta_i; \Omega_{it}^\theta) \propto \exp \left\{ - \sum_{t=1}^{T} \left( \frac{\Omega_{i}^\theta - \theta_i}{2 \sigma_{\omega \theta}^2} \right)^2 \right\} \]

Third, the density kernel of the posterior distribution of capacity is

(A15) \[ L(\theta_i; \Omega_{it}^\theta) \gamma(\theta_i) \propto \frac{\exp \left\{ - \frac{(\theta_i^0 - \theta_i)^2}{2 \sigma_{\omega \theta}^2} \right\} \exp \left\{ - \frac{(\theta_i - \theta_i^e)^2}{2 \sigma_{\theta i}^2} \right\}}{P(\theta_i|0 \leq \theta_i \leq 1, \theta_i^e)} \]

\[ \propto \frac{\frac{1}{2} \exp \left\{ \frac{T}{\sigma_{\omega \theta}^2} \left[ (\Omega_i^\theta - \theta_i^0)^2 + (\Omega_i^\theta - \theta_i^e)^2 \right] \right\} + \frac{1}{\sigma_{\theta i}^2} \theta_i^2}{P(\theta_i|0 \leq \theta_i \leq 1, \theta_i^e)} \]

\[ \propto \frac{\theta_i^2 + \sum_{t=1}^{T} \frac{1}{\sigma_{\theta \theta}^2} \theta_i^2 + \frac{1}{\sigma_{\theta \theta}^2} \theta_i}{\sum_{t=1}^{T} \frac{1}{\sigma_{\theta \theta}^2} \theta_i + \frac{1}{\sigma_{\theta \theta}^2} \theta_i} \]

Completing the square in the numerator of (A15), we have
\[
L(\theta; \Omega^\theta_i) \propto \frac{\exp \left\{ -\frac{1}{2} \left[ \frac{1}{\sigma^2_{\omega \theta} \sigma^2_{\theta e}} \left[ \theta_i - \left( \frac{\sigma^2_{\omega \theta} \sigma^2_{\theta e}}{\sigma^2_{\omega \theta} + T \sigma^2_{\theta e}} \left( \frac{1}{\sigma^2_{\theta e}} \theta_i + \frac{T}{\sigma^2_{\omega \theta}} \Omega^\theta \right) \right] \right] \right\}}{\sqrt{\sigma^2_{\omega \theta} + T \sigma^2_{\theta e}}}
\]

(A16) 

where \( P(\theta_i | 0 \leq \theta_i \leq 1, \theta^e_i) = \Phi \left( \frac{1 - \theta^e_i}{\sigma^2_{\theta e}} \right) - \Phi \left( \frac{-\theta^e_i}{\sigma^2_{\theta e}} \right) \). Therefore, this density kernel is still a truncated normal distribution, in which the posterior variance of capacity is

(A17) 

\[
\Sigma(\theta_i) = \frac{\sigma^2_{\omega \theta} \sigma^2_{\theta e}}{\sigma^2_{\omega \theta} + T \sigma^2_{\theta e}}
\]

and the posterior mean of capacity is

(A18) 

\[
S(\theta_i) = \Sigma(\theta_i) \left[ \frac{1}{\sigma^2_{\theta e}} \theta_i^e + \frac{T}{\sigma^2_{\omega \theta}} \Omega^\theta \right] = \frac{\sigma^2_{\omega \theta}}{\sigma^2_{\omega \theta} + T \sigma^2_{\theta e}} \theta_i^e + \frac{T \sigma^2_{\theta e}}{\sigma^2_{\omega \theta} + T \sigma^2_{\theta e}} \Omega^\theta
\]

As \( \Omega^\theta_i \sim N(\theta_i, \sigma^2_{\omega \theta}) \) and \( T = T\. \text{End} \rightarrow +\infty \), we have

(A19) 

\[
\Omega^\theta_i = \frac{1}{T} \sum_{t=1}^{T} \Omega^\theta_{it} = E(\Omega^\theta_{it}) = \theta_i
\]

thus, the posterior mean of capacity is

(A20) 

\[
S(\theta_i) = \frac{\sigma^2_{\omega \theta}}{\sigma^2_{\omega \theta} + T \sigma^2_{\theta e}} \theta_i^e + \frac{T \sigma^2_{\theta e}}{\sigma^2_{\omega \theta} + T \sigma^2_{\theta e}} \theta_i
\]
B. Comparison of the Expectations and Variances of the Composite Virtue-and-capacity of Village Leaders Before and After the Introduction of Local Direct Elections

B.1 Expectation: We compare, in a representative village, the expectation of the composite virtue-and-capacity of the elected village leader and the appointed village leader. Specifically, we compare

\[
\mathbb{E}^{\text{Ele}}(\mu\alpha_i + (1 - \mu)\theta_i) = \int_0^1 \frac{f_0^1 [\mu \alpha_i + (1 - \mu)\theta_i] A_i^{\text{Ele}} d\alpha_i d\theta_i}{f_0^1 A_i^{\text{Ele}} d\alpha_i d\theta_i}
\]

and

\[
\mathbb{E}^{\text{App}}(\mu\alpha_i + (1 - \mu)\theta_i) = \int_0^1 \frac{f_0^1 [\mu \alpha_i + (1 - \mu)\theta_i] A_i^{\text{App}} d\alpha_i d\theta_i}{f_0^1 A_i^{\text{App}} d\alpha_i d\theta_i}
\]

and we know that

\[
A_i^{\text{Ele}}(\alpha) = \mu S^{\text{Ele}}(\alpha_i) + (1 - \mu) S^{\text{Ele}}(\theta_i)
\]

where

\[
S^{\text{Ele}}(\alpha_i) = \frac{\sigma^{\text{Ele}}(\alpha_i)}{\sigma^{\text{e}_1}} \alpha_i + \left(1 - \frac{\sigma^{\text{Ele}}(\alpha_i)}{\sigma^{\text{e}_1}}\right) \alpha_i
\]

\[
= \frac{\sigma^{\text{e}_2}}{\sigma^{\text{e}_1}} \alpha_i + \frac{\sigma^{\text{e}_1}}{\sigma^{\text{e}_1}} \alpha_i
\]

\[
S^{\text{Ele}}(\theta_i) = \left[\frac{\sigma^{\text{Ele}}(\theta_i)}{\sigma^{\text{e}_1}} \theta_i + \left(1 - \frac{\sigma^{\text{Ele}}(\theta_i)}{\sigma^{\text{e}_1}}\right) \theta_i\right]
\]

As \( T^{\text{Ele}} > T^{\text{App}} \), comparing the composite virtue-and-capacity of the elected village leader and appointed village leader is equivalent to having

\[
f(x(T), y(T)) = \frac{f_0^1 f_0^1 [\mu \alpha_i + (1 - \mu)\theta_i] S^{\text{Ele}}(\alpha_i) + (1 - \mu) S^{\text{Ele}}(\theta_i)] d\alpha_i d\theta_i}{f_0^1 f_0^1 S^{\text{Ele}}(\alpha_i) + (1 - \mu) S^{\text{Ele}}(\theta_i)] d\alpha_i d\theta_i}
\]

where \( S(\alpha_i) = (1 - x) \alpha_i^e + x \alpha_i \), \( x = \frac{\sigma^{\text{e}_1} T}{\sigma^{\text{e}_1} + \sigma^{\text{e}_1}} \); \( S(\theta_i) = (1 - y) \theta_i^e + y \theta_i \), \( y = \frac{\sigma^{\text{e}_1} T}{\sigma^{\text{e}_1} + \sigma^{\text{e}_1}} \). Then we need to prove that

\[
\frac{\sigma_x^T}{\sigma^{\text{e}_1} + \sigma^{\text{e}_1}} \cdot T = T^{\text{Ele}} or T^{\text{App}}. Then we need to prove that
\]

\[
\frac{\partial f}{\partial T} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial T} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial T} > 0
\]

It is easy to prove that \( \frac{\partial x}{\partial T} > 0 \) and \( \frac{\partial y}{\partial T} > 0 \), so we only need to prove that \( \frac{\partial f}{\partial x} \geq 0 \) and \( \frac{\partial f}{\partial y} \geq 0 \).

First, simplify the numerator of (B6)
\[
\int_0^1 \int_0^1 [\mu \alpha_i + (1 - \mu) \theta_i] \{ \mu l[(1 - x) \alpha^e_i + x \alpha_i] + (1 - \mu) l[(1 - y) \theta^e_i + y \theta_i] \} \, d \alpha_i \, d \theta_i
\]

\[
= \int_0^1 \int_0^1 \left\{ \mu^2 l \alpha_i [(1 - x) \alpha^e_i + x \alpha_i] + (1 - \mu)^2 l \theta_i [(1 - y) \theta^e_i + y \theta_i] + \mu (1 - \mu) l \alpha_i [(1 - x) \alpha^e_i + x \alpha_i] + \mu (1 - \mu) l \theta_i [(1 - x) \alpha^e_i + x \alpha_i] \right\} \, d \alpha_i \, d \theta_i
\]

where

(B8) \[
\int_0^1 \int_0^1 \mu^2 l \alpha_i [(1 - x) \alpha^e_i + x \alpha_i] \, d \alpha_i \, d \theta_i
\]

\[
= \int_0^1 \int_0^1 \mu^2 l(1 - x) \alpha^e_i \alpha_i \, d \alpha_i \, d \theta_i + \int_0^1 \int_0^1 \mu^2 l x \alpha^2_i \, d \alpha_i \, d \theta_i
\]

\[
= \frac{1}{2} \mu^2 l(1 - x) \alpha^e_i + \frac{1}{3} \mu^2 l x
\]

(B9) \[
\int_0^1 \int_0^1 (1 - \mu)^2 l[(1 - y) \theta^e_i + y \theta_i] \, d \alpha_i \, d \theta_i
\]

\[
= \int_0^1 \int_0^1 (1 - \mu)^2 l(1 - y) \theta^e_i \theta_i \, d \alpha_i \, d \theta_i + \int_0^1 \int_0^1 (1 - \mu)^2 l y \theta^2_i \, d \alpha_i \, d \theta_i
\]

\[
= \frac{1}{2} (1 - \mu)^2 l(1 - y) \theta^e_i + \frac{1}{3} (1 - \mu)^2 l y
\]

(B10) \[
\int_0^1 \int_0^1 \mu (1 - \mu) l \alpha_i [(1 - y) \theta^e_i + y \theta_i] \, d \alpha_i \, d \theta_i
\]

\[
= \int_0^1 \int_0^1 \mu (1 - \mu) l \alpha_i (1 - y) \theta^e_i \, d \alpha_i \, d \theta_i + \int_0^1 \int_0^1 \mu (1 - \mu) l \alpha_i y \theta_i \, d \alpha_i \, d \theta_i
\]

\[
= \frac{1}{2} \mu (1 - \mu)(1 - y) \theta^e_i + \frac{1}{4} \mu (1 - \mu) l y
\]

(B11) \[
\int_0^1 \int_0^1 \mu (1 - \mu) l \theta_i [(1 - x) \alpha^e_i + x \alpha_i] \, d \alpha_i \, d \theta_i
\]

\[
= \int_0^1 \int_0^1 \mu (1 - \mu) l (1 - x) \alpha^e_i \, d \alpha_i \, d \theta_i + \int_0^1 \int_0^1 \mu (1 - \mu) l x \alpha_i \, d \alpha_i \, d \theta_i
\]

\[
= \frac{1}{2} \mu (1 - \mu) l (1 - x) \alpha^e_i + \frac{1}{4} \mu (1 - \mu) l x
\]

Therefore, the numerator of (B6) becomes

(B12) \[
\frac{1}{2} \mu l(1 - x) \alpha^e_i + \frac{1}{2} (1 - \mu) l(1 - y) \theta^e_i + \frac{1}{12} \mu (\mu + 3) l x + \frac{1}{12} (\mu - 4)(\mu - 1) l y
\]

Second, simplify the denominator of (B6)

(B13) \[
\int_0^1 \int_0^1 \mu l[(1 - x) \alpha^e_i + x \alpha_i] + (1 - \mu) l[(1 - y) \theta^e_i + y \theta_i] \, d \alpha_i \, d \theta_i
\]

\[
= \int_0^1 \int_0^1 \mu l(1 - x) \alpha^e_i \, d \alpha_i \, d \theta_i + \int_0^1 \int_0^1 \mu l x \alpha_i \, d \alpha_i \, d \theta_i
\]

\[
+ \int_0^1 \int_0^1 (1 - \mu) l(1 - y) \theta^e_i \, d \alpha_i \, d \theta_i + \int_0^1 \int_0^1 (1 - \mu) l y \theta_i \, d \alpha_i \, d \theta_i
\]

\[
= \mu l(1 - x) \alpha^e_i + \frac{1}{2} \mu l x + (1 - \mu) l(1 - y) \theta^e_i + \frac{1}{2} (1 - \mu) l y
\]

Accordingly, (B6) becomes

(B14) \[
f(x(T), y(T)) = \frac{\frac{1}{2} \mu l(1 - x) \alpha^e_i + \frac{1}{2} (1 - \mu) l(1 - y) \theta^e_i + \frac{1}{12} \mu (\mu + 3) l x + \frac{1}{12} (\mu - 4)(\mu - 1) l y}{\mu l(1 - x) \alpha^e_i + (1 - \mu) l(1 - y) \theta^e_i + \frac{1}{2} \mu l x + \frac{1}{2} (1 - \mu) l y}
\]

where the first order derivative of the numerator of (B14) with respect to x is

\[-\frac{1}{2} \mu \alpha^e_i + \frac{1}{12} \mu (\mu + 3) l , \text{and the first order derivative of the denominator of (B14) with}

\]
respect to $x$ is $-\mu \alpha_i^e + \frac{1}{2} \mu l$. 

Then, we can obtain 

(B15) \[ \frac{\partial f}{\partial x} = \frac{\frac{1}{12} \mu^2 \theta_i^e + \frac{1}{12} \mu (1-\mu)^2 (1-x) \theta_i^e + \frac{1}{12} \mu^2 (1-\mu)^2 (1-x) \theta_i^e + \frac{1}{2} \mu (1-\mu) (1-2 \mu l) l x }{[\mu (1-x) \alpha_i^e + (1-\mu) (1-\mu) \theta_i^e + \frac{1}{2} \mu (1-\mu) l x + \frac{1}{2} (1-\mu) l x]^2} \]

To ensure that $\frac{\partial f}{\partial x} \geq 0$, we need to have $\mu (1-\mu) (2\mu - 1) \geq 0$, as shown in the first step of (B15), that is, $1 \geq \mu \geq \frac{1}{2}$ or $\mu = 0$. However, when $0 < \mu < \frac{1}{2}$, we need to rearrange the items of the numerator in (B15), as shown in the second step of (B15), so that when we have $2\mu (1-\alpha_i^e) + 2\alpha_i^e - 1 \geq 2\mu (1-\alpha_i^e) \geq 0$, that is, $\alpha_i^e \geq \frac{1}{2}$, we have $\frac{\partial f}{\partial x} \geq 0$. In other words, $\frac{\partial f}{\partial x} \geq 0$ holds, when $1 \geq \mu \geq \frac{1}{2}$ or $\mu = 0$, or both $0 < \mu < \frac{1}{2}$ and $\alpha_i^e \geq \frac{1}{2}$.

Now we discuss $\frac{\partial f}{\partial y}$. The first order derivative of the numerator of (B14) with respect to $y$ is \[-\frac{1}{2} (1-\mu) l \cdot \theta_i^e + \frac{1}{12} (\mu - 4) (\mu - 1) l,\] and the first order derivative of the denominator of (B14) with respect to $y$ is \[-(1-\mu) l \cdot \theta_i^e + \frac{1}{2} (1-\mu) l.\]

Then, we can obtain 

(B16) \[ \frac{\partial f}{\partial y} = \frac{\frac{1}{12} \mu^2 (1-\mu) l x \theta_i^e + \frac{1}{12} (1-\mu)^3 \theta_i^e + \frac{1}{12} \mu (1-\mu)^2 l (1-x) \alpha_i^e + \frac{1}{2} \mu (1-\mu) (1-2 \mu l) l x }{[\mu (1-x) \alpha_i^e + (1-\mu) (1-\mu) \theta_i^e + \frac{1}{2} \mu (1-\mu) l x + \frac{1}{2} (1-\mu) l x]^2} \]

To ensure that $\frac{\partial f}{\partial y} \geq 0$, we need to have $\mu (1-\mu) (1-2 \mu) \geq 0$, as shown in the first step of (B16), that is, $0 \leq \mu \leq \frac{1}{2}$ or $\mu = 1$. However, when $\frac{1}{2} < \mu < 1$, we need to rearrange the items of the numerator in (B16), as shown in the second step of (B16), so that when we have $1 - 2 \mu + 2 \mu \cdot \theta_i^e \geq 1 - 2 \mu + u \geq 0$, that is, $\theta_i^e \geq \frac{1}{2}$, we have $\frac{\partial f}{\partial y} \geq 0$. In other words, $\frac{\partial f}{\partial y} \geq 0$ holds, when $0 \leq \mu \leq \frac{1}{2}$ or $\mu = 1$, or both $\frac{1}{2} < \mu < 1$ and $\theta_i^e \geq \frac{1}{2}$.

By proving that $\frac{\partial f}{\partial x} \geq 0$ and $\frac{\partial f}{\partial y} \geq 0$, we prove that $\frac{\partial f}{\partial T} > 0$. As $T^{ELe} > T^{App}$, we know that $\mathbb{E}^{ELe} (\mu \alpha_i + (1-\mu) \theta_i) > \mathbb{E}^{App} (\mu \alpha_i + (1-\mu) \theta_i)$. 

43
In summary, when \( \mu = 0, 1, \frac{1}{2} \), we have \( \mathbb{E}^{E} (\mu \alpha_i + (1 - \mu) \theta_i) > \mathbb{E}^{App} (\mu \alpha_i + (1 - \mu) \theta_i) \). When \( 0 < \mu < \frac{1}{2} \), by assuming that \( \alpha_i^e \geq \frac{1}{2} \), we have \( \mathbb{E}^{E} (\mu \alpha_i + (1 - \mu) \theta_i) > \mathbb{E}^{App} (\mu \alpha_i + (1 - \mu) \theta_i) \). When \( \frac{1}{2} < \mu < 1 \), by assuming that \( \theta_i^e \geq \frac{1}{2} \), we have \( \mathbb{E}^{E} (\mu \alpha_i + (1 - \mu) \theta_i) > \mathbb{E}^{App} (\mu \alpha_i + (1 - \mu) \theta_i) \).

B.2 Variance: We then compare, in a representative village, the variance of the composite virtue-and-capacity of the elected village leader and appointed village leader.

We know that

\[
\begin{align*}
\text{(B17)} & \quad \text{Var}^{E} (\mu \alpha_i + (1 - \mu) \theta_i) \\
& = \mathbb{E}^{E} ((\mu \alpha_i + (1 - \mu) \theta_i)^2) - \mathbb{E}^{E} (\mu \alpha_i + (1 - \mu) \theta_i)^2
\end{align*}
\]

We first define \( g(x(T), y(T)) \equiv \mathbb{E}^{E} ((\mu \alpha_i + (1 - \mu) \theta_i)^2) \) and we have

\[
\begin{align*}
\text{(B18)} & \quad g(x(T), y(T)) = \frac{\int_0^1 \int_0^1 \mu^2 \alpha_i^2 + (1 - \mu) \theta_i^2 + 2 \mu(1 - \mu) \mu \alpha_i \theta_i}{\int_0^1 \int_0^1 \mu^2 (\alpha_i + 1 - \mu) \alpha_i \theta_i}
\end{align*}
\]

where \( S(\alpha_i) = (1 - x) \alpha_i^e + x \alpha_i, \ S(\theta_i) = (1 - y) \theta_i^e + y \theta_i, \ x \equiv \frac{\sigma^2_{\theta \xi T}}{\sigma^2_{\theta \alpha} + \sigma^2_{\theta e T}}, \) and \( y \equiv \frac{\sigma^2_{\theta \xi T}}{\sigma^2_{\theta \alpha} + \sigma^2_{\theta e T}} \).

First, we simplify the numerator of \( g(\cdot) \)

\[
\begin{align*}
\text{(B19)} & \quad \int_0^1 \int_0^1 \mu^2 \alpha_i^2 + (1 - \mu)^2 \theta_i^2 + 2 \mu(1 - \mu) \mu \alpha_i \theta_i, \ \text{d} \alpha_i \text{d} \theta_i \\
& = \int_0^1 \int_0^1 \mu^2 l(1 - \mu) \alpha_i^2 [(1 - x) \alpha_i^e + x \alpha_i] + (1 - \mu)^3 \theta_i^2 [(1 - y) \theta_i^e + y \theta_i]
\end{align*}
\]

where

\[
\begin{align*}
\text{(B20)} & \quad \int_0^1 \int_0^1 \mu^3 \alpha_i^2 [(1 - x) \alpha_i^e + x \alpha_i] \text{d} \alpha_i \text{d} \theta_i = \frac{1}{3} \mu^3 l(1 - x) \alpha_i^e + \frac{1}{4} \mu^3 lx \\
\text{(B21)} & \quad \int_0^1 \int_0^1 \mu^2 l(1 - \mu) \alpha_i^2 [(1 - y) \theta_i^e + y \theta_i] \text{d} \alpha_i \text{d} \theta_i \\
& = \frac{1}{3} \mu^2 l(1 - \mu)(1 - y) \theta_i^e + \frac{1}{6} \mu^2 l(1 - \mu)y \\
\text{(B22)} & \quad \int_0^1 \int_0^1 \mu(1 - \mu)^2 \theta_i^2 [(1 - x) \alpha_i^e + x \alpha_i] \text{d} \alpha_i \text{d} \theta_i \\
& = \frac{1}{3} \mu(1 - \mu)^2 l(1 - x) \alpha_i^e + \frac{1}{6} \mu l(1 - \mu) x \\
\text{(B23)} & \quad \int_0^1 \int_0^1 (1 - \mu)^3 \theta_i^2 [(1 - y) \theta_i^e + y \theta_i] \text{d} \alpha_i \text{d} \theta_i \\
& = \frac{1}{3} (1 - \mu)^3 l(1 - y) \theta_i^e + \frac{1}{4} (1 - \mu)^3 ly
\end{align*}
\]
\[ \int_0^1 \int_0^1 2\mu^2 l(1-\mu)\alpha_i \theta_i [(1-x)\alpha_i^e + x\alpha_i^m] \, d\alpha_i \, d\theta_i = \frac{1}{2} \mu^2 (1-\mu) l(1-x)\alpha_i^e + \frac{1}{3} \mu^2 (1-\mu)lx \]

(B24) \[
\int_0^1 \int_0^1 2\mu(1-\mu)^2 \alpha_i \theta_i [(1-y)\theta_i^e + y\theta_i^m] \, d\alpha_i \, d\theta_i = \frac{1}{2} \mu (1-\mu)^2 l(1-y)\theta_i^e + \frac{1}{2} \mu (1-\mu)^2ly 
\]

The numerator of \( g(\cdot) \) becomes \( \frac{1}{6} \mu (\mu^2 - \mu + 2) l(1-x)\alpha_i^e + \frac{1}{6} (1-\mu)(\mu^2 - \mu + 2) l(1-y)\theta_i^e + \frac{1}{12} \mu (\mu^2 + 2)lx + \frac{1}{12} (1-\mu)(\mu^2 - 2\mu + 3)ly \). Therefore, we have

(B26) \[
g(\cdot) = \frac{\frac{1}{2} \mu (\mu^2 - \mu + 2) l(1-x)\alpha_i^e + \frac{1}{6} (1-\mu)(\mu^2 - \mu + 2) l(1-y)\theta_i^e + \frac{1}{12} \mu (\mu^2 + 2)lx + \frac{1}{12} (1-\mu)(\mu^2 - 2\mu + 3)ly}{\mu l(1-x)\alpha_i^e + (1-\mu) l(1-y)\theta_i^e + \frac{1}{2} \mu lx + \frac{1}{2} (1-\mu)ly} 
\]

Then, we define \( f(x(T), y(T)) \equiv E^{Eie} (\text{App}) [(\mu\alpha_i + (1 - \mu)\theta_i)] \) and we know that

\[
f(x(T), y(T)) = \frac{\frac{1}{2} l(1-x)\alpha_i^e + \frac{1}{6} (1-\mu)(\mu^2 + 3) ly + \frac{1}{12} (\mu - 1) l(1-\mu)ly}{\mu l(1-x)\alpha_i^e + (1-\mu) l(1-y)\theta_i^e + \frac{1}{2} \mu lx + \frac{1}{2} (1-\mu)ly} \]

from (B14).

From (B17), we know that

(B27) \[
\frac{\partial \text{Var}(\cdot)}{\partial T} = \frac{\partial g(\cdot)}{\partial T} = 2f(\cdot) \cdot \frac{\partial f(\cdot)}{\partial T} \]

It is known that \( f(\cdot) > 0, \frac{\partial f(\cdot)}{\partial T} > 0 \). We then need to derive \( \frac{\partial g(\cdot)}{\partial T} = \frac{\partial g(\cdot)}{\partial x} \cdot \frac{\partial x}{\partial T} + \frac{\partial g(\cdot)}{\partial y} \cdot \frac{\partial y}{\partial T} \).

\[
\frac{\partial y}{\partial T} \quad \text{It is known that} \quad 0 < \frac{\partial x}{\partial T} > 0, \text{therefore, we have}
\]

(B28) \[
\frac{\partial g(\cdot)}{\partial x} = \frac{\frac{1}{2} l^2 (1-x)\alpha_i^e + \frac{1}{12} l^2 (1-\mu)(\mu^2 + 2) l(1-y)\theta_i^e + \mu l(1-x)\alpha_i^e + \frac{1}{12} l^2 (1-\mu)(\mu^2 + 2) l(1-y)\theta_i^e + \mu l(1-\mu)lx + \frac{1}{2} (1-\mu)ly}{\mu l(1-x)\alpha_i^e + (1-\mu) l(1-y)\theta_i^e + \frac{1}{2} \mu lx + \frac{1}{2} (1-\mu)ly} 
\]

(B29) \[
\frac{\partial g(\cdot)}{\partial y} = \frac{\frac{1}{2} l^2 (1-x)\alpha_i^e + \frac{1}{12} l^2 (1-\mu)(\mu^2 + 2) l(1-y)\theta_i^e + \mu l(1-x)\alpha_i^e + \frac{1}{12} l^2 (1-\mu)(\mu^2 + 2) l(1-y)\theta_i^e + \mu l(1-\mu)lx + \frac{1}{2} (1-\mu)ly}{\mu l(1-x)\alpha_i^e + (1-\mu) l(1-y)\theta_i^e + \frac{1}{2} \mu lx + \frac{1}{2} (1-\mu)ly} 
\]

Thus,

\[
\frac{\partial g(\cdot)}{\partial T} = \frac{\frac{1}{2} l^2 (1-x)\alpha_i^e + \frac{1}{12} l^2 (1-\mu)(\mu^2 + 2) l(1-y)\theta_i^e + \mu l(1-x)\alpha_i^e + \frac{1}{12} l^2 (1-\mu)(\mu^2 + 2) l(1-y)\theta_i^e + \mu l(1-\mu)lx + \frac{1}{2} (1-\mu)ly}{\mu l(1-x)\alpha_i^e + (1-\mu) l(1-y)\theta_i^e + \frac{1}{2} \mu lx + \frac{1}{2} (1-\mu)ly} \cdot \frac{\partial x}{\partial T} 
\]

\[
+ \frac{1}{2} l^2 (1-x)\alpha_i^e + \frac{1}{12} l^2 (1-\mu)(\mu^2 + 2) l(1-y)\theta_i^e + \mu l(1-x)\alpha_i^e + \frac{1}{12} l^2 (1-\mu)(\mu^2 + 2) l(1-y)\theta_i^e + \mu l(1-\mu)lx + \frac{1}{2} (1-\mu)ly}{\mu l(1-x)\alpha_i^e + (1-\mu) l(1-y)\theta_i^e + \frac{1}{2} \mu lx + \frac{1}{2} (1-\mu)ly} \cdot \frac{\partial y}{\partial T} 
\]

It is easy to prove that \( \frac{\partial g(\cdot)}{\partial T} = \frac{\partial f(\cdot)}{\partial T} \). In addition, from (B14), we can deduce that \( f(\cdot) > \frac{1}{2} \) as the numerator of \( f(\cdot) \) is twice as large as the denominator of \( f(\cdot) \). As a result, we can deduce from (B27) that \( \frac{\partial \text{Var}(\cdot)}{\partial T} = \frac{\partial g(\cdot)}{\partial T} - 2f(\cdot) \cdot \frac{\partial f(\cdot)}{\partial T} < 0 \). As \( T^{Eie} > T^{App} \), we know that \( \text{Var}^{Eie} (\mu\alpha_i + (1 - \mu)\theta_i) < \text{Var}^{App} (\mu\alpha_i + (1 - \mu)\theta_i) \).
C. Posterior Means of the Composite Virtue-and-Capacity of Village Party Secretary Candidates

We know that for the composite virtue-and-capacity of village party committee members, the distribution of the prior perceptions of the representative township official is \( N(\mu \alpha^u + (1 - \mu)\theta^u, \mu^2 \sigma^2_u + (1 - \mu)^2 \sigma^2_\theta u) \). By observing the performance in a series of tasks \( P_{jm} = \mu \alpha_j + (1 - \mu)\theta_j + \varepsilon_{jm} \), where \( m = 1, 2, ..., M \), with \( M \) times of iterations, we can obtain the posterior distribution of the composite virtue-and-capacity of village party committee member \( j \). For simplicity, note that \( \pi^u \equiv \mu \alpha^u + (1 - \mu)\theta^u \), \( \pi_j \equiv \mu \alpha_j + (1 - \mu)\theta_j \), and \( \sigma^2_u \equiv \mu^2 \sigma^2_u + (1 - \mu)^2 \sigma^2_\theta u \).

Following Koop et al. (2007), we derive the density kernel of the posterior distribution of the composite virtue-and-capacity, such that

\[
\Pr(\pi_j | P_{jm}) = \gamma(\pi_j) \cdot [L(\pi_j; P_{j1}) \ldots L(\pi_j; P_{j2}) \ldots L(\pi_j; P_{jM})]
\]

where the vector \( \Omega^u_{it} = [\Omega^u_{i1} \ldots \Omega^u_{it} \ldots \Omega^u_{it}] \).

First, we derive \( \gamma(\pi_j) \), the density kernel of the prior distribution of the composite virtue-and-capacity. As the prior distribution is a normal distribution truncated at \( [0, 1] \), we have

\[
\gamma(\pi_j) = \gamma(\pi_j | 0 \leq \pi_j \leq 1, \pi^u)
\]

\[
\propto \exp\left(\frac{-(\pi_j - \pi^u)^2}{2\sigma^2_u}\right)
\]

where \( \Pr(\pi_j | 0 \leq \pi_j \leq 1, \pi^u) = \Phi\left(\frac{1-\pi^u}{\sigma^2_u}\right) - \Phi\left(\frac{-\pi^u}{\sigma^2_u}\right) \) represents the probability that \( \pi_j \) is at \( [0, 1] \), contingent on \( \pi^u \).

Second, we derive the density of \( P_{jm} \) as a likelihood function given by

\[
L(\pi_j; P_{jm}) \propto \exp\left(-M \frac{(P_{jm} - \pi^u)^2}{2\sigma^2}\right)
\]

Third, by Bayes’ rule, the density kernel of the posterior distribution of the composite virtue-and-capacity of village party committee member \( j \) is

\[
\Pr(\pi_j; P_{jm}) \cdot \gamma(\pi_j) \propto \frac{\exp\left(-\frac{1}{2} \frac{1}{\sigma^2_u} | \pi_j - \left(\frac{\sigma^2_u}{\sigma^2_u + M\sigma^2_u} P_{jm} - \frac{\sigma^2_u}{\sigma^2_u + M\sigma^2_u} | \pi^u \right|^{2}\right)}{\Pr(\pi_j | 0 \leq \pi_j \leq 1, \pi^u)}
\]

where \( \Pr(\pi_j | 0 \leq \pi_j \leq 1, \pi^u) = \Phi\left(\frac{1-\pi^u}{\sigma^2_u}\right) - \Phi\left(\frac{-\pi^u}{\sigma^2_u}\right) \). Therefore, this density kernel is still a truncated normal distribution, and the posterior mean of the composite virtue-and-capacity of village party committee member \( j \) is
\[(C5) \quad S(\pi_j) = \frac{\sigma^2 \pi \sum_j^M P_j}{\sigma^2 + M \sigma_u^2} + \frac{M \pi_j}{\sigma^2} = \frac{\sigma^2 \pi}{\sigma^2 + M \sigma_u^2} + \frac{M \sigma_u^2}{\sigma^2 + M \sigma_u^2} \bar{P}_j\]

As \( P_{jm} \sim N(\pi_j, \sigma^2) \) and \( M \to + \infty \), we have

\[(C6) \quad \bar{P}_j = \frac{1}{M} \sum_{m=1}^M P_{jm} = E(P_{jm}) = \pi_j\]

thus, the posterior mean of the composite virtue-and-capacity of village party committee member \( j \) is

\[(C7) \quad S(\pi_j) = \frac{\sigma^2 \pi}{\sigma^2 + M \sigma_u^2} + \frac{M \sigma_u^2}{\sigma^2 + M \sigma_u^2} \pi_j\]
D. Comparison of the Expectations of the Composite Virtue-and-Capacity of Village Party Secretaries Before and After the Introduction of Local Direct Elections

We compare, in a representative village, the expectation of the composite virtue-and-capacity of the village party secretary before and after the introduction of local direct elections. Note that \( \pi_{jk} \equiv \mu \alpha_{jk} + (1 - \mu) \theta_{jk} \), \( \pi_{k} \equiv \mu \alpha_{k} + (1 - \mu) \theta_{k} \), and \( \pi^{u} \equiv \mu \alpha^{u} + (1 - \mu) \theta^{u} \). Specifically, we compare

\[
\[\mu \alpha + (1 - \mu) \theta\]^{\text{VPS,After}} = \frac{\int_{0}^{1} [\pi_{jk}]^{\text{Ele}} R_{jk}^{\text{Ele}} d\pi_{jk} + \int_{0}^{1} \pi_{jk} R_{jk}^{\text{Ele}} d\pi_{jk}}{\int_{0}^{1} R_{jk}^{\text{Ele}} d\pi_{jk} + \int_{0}^{1} R_{jk}^{\text{Ele}} d\pi_{jk}}
\]

and

\[
\[\mu \alpha + (1 - \mu) \theta\]^{\text{VPS,Before}} = \frac{\int_{0}^{1} [\pi_{jk}]^{\text{App}} R_{jk}^{\text{App}} d\pi_{jk} + \int_{0}^{1} \pi_{jk} R_{jk}^{\text{App}} d\pi_{jk}}{\int_{0}^{1} R_{jk}^{\text{App}} d\pi_{jk} + \int_{0}^{1} R_{jk}^{\text{App}} d\pi_{jk}}
\]

and we know that

\[
R_{jk}^{\text{Ele (App)}} = (1 - \varphi) w^{u} + \varphi w [\pi_{jk}]^{\text{Ele (App)}}
\]

and

\[
R_{jk} = (1 - \varphi) w^{u} + \varphi w \pi_{jk}
\]

Then we have

\[
\int_{0}^{1} \pi_{jk} d\pi_{jk} \equiv E(\pi_{jk}) \text{ represents the expectation of the composite virtue-and-capacity of all village committee members. As discussed in Sections III.A. and III.B., the inferences and selection of each village committee member, including the village leader as representative, are homogenous. Therefore, the expectation of the composite virtue-and-capacity of each village committee member is homogenous. In other words, } E_{jk}(\pi_{jk}), \text{ the expectation of the composite virtue-and-capacity, is the same for each } j_k.
\]

As a result, \( E(\pi_{jk}) = E(E_{jk}(\pi_{jk})) = E_1(\pi_1) \), where \( j_k = 1 \) represents the village leader as representative of all village committee members, and \( E_1(\pi_1) \) represents the expectation of the composite virtue-and-capacity of the village leader, such that

\[
E_1(\pi_1) \equiv f(x(T), y(T)) = \frac{\int_{0}^{1} \int_{0}^{1} [\mu \alpha_{i} + (1 - \mu) \theta_{i}] [\mu \alpha_{i} + (1 - \mu) \theta_{i}] d\alpha_{i} d\theta_{i}}{\int_{0}^{1} \int_{0}^{1} [\mu \alpha_{i} + (1 - \mu) \theta_{i}] [\mu \alpha_{i} + (1 - \mu) \theta_{i}] d\alpha_{i} d\theta_{i}}
\]

48
where \( S(\alpha_i) = (1-x)\alpha_i^e + x\alpha_i \), \( x \equiv \frac{\sigma^2_{\theta}T}{\sigma^2_{\omega}+\sigma^2_{\theta}eT} \); \( S(\theta_i) = (1-y)\theta_i^e + y\theta_i \), \( y \equiv \frac{\sigma^2_{\theta}T}{\sigma^2_{\omega}+\sigma^2_{\theta}eT} \).

Similarly, \( \int_0^1 \pi^2_{jk} d\pi_{jk} \equiv E(\pi^2_{jk}) \) represents the expectation of the square composite virtue-and-capacity of all village committee members. As discussed in Sections III.A. and III.B., the inferences and selection of each village committee member, including the village leader as representative, are homogenous. Therefore, the expectation of the square composite virtue-and-capacity of each village committee member is homogenous. In other words, \( E_j(\pi^2_{jk}) \), the expectation of the square composite virtue-and-capacity, is the same for each \( j_k \). As a result, \( E(\pi^2_{jk}) = E(E_j(\pi^2_{jk})) = E_1(\pi^2_j) \), where \( j_k = 1 \) represents the village leader as representative of all village committee members, and \( E_1(\pi^2_j) \) represents the expectation of the square composite virtue-and-capacity of the village leader, such that

\[
E_1(\pi^2_j) \equiv g(x(T), y(T)) = \int_0^1 \int_0^1 \mu \alpha_i + (1-\mu) \theta_i^e = \int_0^1 \int_0^1 \mu \alpha_i + (1-\mu) \theta_i^e \] ![](https://latex.codecogs.com/svg.image?\int_0^1 \int_0^1 \mu \alpha_i + (1-\mu) \theta_i^e \]

where \( S(\alpha_i) = (1-x)\alpha_i^e + x\alpha_i \), \( S(\theta_i) = (1-y)\theta_i^e + y\theta_i \), \( x \equiv \frac{\sigma^2_{\theta}T}{\sigma^2_{\omega}+\sigma^2_{\theta}eT} \), and \( y \equiv \frac{\sigma^2_{\theta}T}{\sigma^2_{\omega}+\sigma^2_{\theta}eT} \).

Therefore, (D5) becomes

\[
\frac{(1-\varphi)w\pi^u f(t)+\varphi g(t)+\frac{1}{2}(1-\varphi)w\pi^u + \frac{1}{3}w^2}{(1-\varphi)w\pi^u + \varphi g(t)+(1-\varphi)w\pi^u + \frac{1}{2}w}.
\]

It is easy to deduce that \( f(T) > g(T) \) and \( f'(T) = g'(T) \). We also know that \( (1-\varphi)w\pi^u + \frac{1}{2}\varphi w > \frac{1}{2}(1-\varphi)w\pi^u + \frac{1}{3}w \). Therefore, the first order derivative of (D8) with respect to \( T \) is greater than 0.

We know that

\[
[\mu \alpha + (1-\mu) \theta]^{VPS, After} = \frac{(1-\varphi)w\pi^u f(T^{Ele})+\varphi g(T^{Ele})+\frac{1}{2}(1-\varphi)w\pi^u + \frac{1}{3}w}{(1-\varphi)w\pi^u + \varphi g(T^{Ele})+(1-\varphi)w\pi^u + \frac{1}{2}w}.
\]

and that

\[
[\mu \alpha + (1-\mu) \theta]^{VPS, Before} = \frac{(1-\varphi)w\pi^u f(T^{App})+\varphi g(T^{App})+\frac{1}{2}(1-\varphi)w\pi^u + \frac{1}{3}w}{(1-\varphi)w\pi^u + \varphi g(T^{App})+(1-\varphi)w\pi^u + \frac{1}{2}w}.
\]

Because \( T^{Ele} > T^{App} \), we know that \( [\mu \alpha + (1-\mu) \theta]^{VPS, After} > [\mu \alpha + (1-\mu) \theta]^{VPS, Before} \).

49