

PDR

The Pakistan Development Review

ISSN 0030 - 9729

Volume 60 - Number 2 - 2021

Editor

Nadeem Ul Haque

Co-editor

Durr-e-Nayab

Associate Editors

Omer Siddique

Naseem Faraz

Literary Editor

Mian Imran Ul Haq

Advisory Board

Ali Cheema

Andrew Barkley

Ashfaq Hasan Khan

Ather Maqsood Ahmed

Deirdre Nansen McCloskey

Diane Coyle

Gavin W. Jones

Ghulam Muhammad Arif

Hafeez Pasha

Hamza Ali Malik

Maleeha Lodhi

Masood Ahmed

Masooda Bano

Matthew McCartney

Pervez Tahir

Shahid Hafeez Kardar

Surjit Bhalla

Syed Akbar Zaidi

Takashi Kurosaki

Articles

Exchange Rate, Output and Macroeconomic Policy:
A Structural Approach

Size and Age as Determinants of Employment Growth
among Manufacturing Firms in Pakistan

Acemoglu, Johnson and Robinson's Notion of Exogenous
Imposition of Colonial Institutions onto Colonies—
A Critique in the Light of Historical Evidence

Friendly Fire: Wheat Subsidy in Punjab, Pakistan

Remittances and Healthcare Expenditures:
Evidence from Pakistan

Commentary

Machine Learning for Economists: An Introduction

Why do We Have Less Investment than China and India?

Policy

The Pension Bomb and Possible Solutions

Books

Looking Back: How Pakistan Became an
Asian Tiger by 2050

C O N T E N T S

Pages

ARTICLES

- Moumita Basu, Ranjanendra Narayan Nag, and Bhaskar Goswami**
Exchange Rate, Output and Macroeconomic Policy: A Structural Approach 103
- Farrukh Iqbal and Aadil Nakhoda**
Size and Age as Determinants of Employment Growth among Manufacturing Firms in Pakistan 119
- Madeeha Gohar Qureshi, Unbreen Qayyum, Musleh ud Din, and Ejaz Ghani**
Acemoglu, Johnson and Robinson's Notion of Exogenous Imposition of Colonial Institutions onto Colonies—A Critique in the Light of Historical Evidence 133
- Muhammad Ahsan Rana and Muhammad Nadeem Malik**
Friendly Fire: Wheat Subsidy in Punjab, Pakistan 153
- Karim Khan, Muhammad Jehangir Khan, and Abid Hussain**
Remittances and Healthcare Expenditures: Evidence from Pakistan 175

COMMENTARY

- Sonan Memon**
Machine Learning for Economists: An Introduction 201
- Hafsa Hina**
Why do We Have Less Investment than China and India? 213

POLICY

- Mahmood Khalid, Naseem Faraz, and Brig. (Retd.) Muhammad Ashraf**
The Pension Bomb and Possible Solutions 225

BOOKS

- Nadeem Ul Haque**
Looking Back: How Pakistan Became an Asian Tiger by 2050
Reviewed by Muhammad Haroon Hussain and Abedullah 231

Exchange Rate, Output and Macroeconomic Policy: A Structuralist Approach

MOUMITA BASU, RANJANENDRA NARAYAN NAG, and BHASKAR GOSWAMI

Current account imbalance and concomitant macroeconomic instability in emerging market economies have been major issues of recent macroeconomic modelling. This paper addresses these issues by asking how international interdependence has impinged on key macroeconomic variables and policy options. There are three assets: domestic bonds, foreign bonds and money. Domestic bonds and foreign bonds are imperfect substitutes due to presence of risk premium. The striking features of the model include endogenous risk premium and balance sheet effect on investment demand due to exchange rate depreciation. We use a simple open economy structuralist macro model that explains the interaction between current account adjustment and exchange rate dynamics. The balance sheet effect and the risk premium together explain how fiscal expansion or monetary expansion may have both short run and long run contractionary effect on the output level with worsening current account balance in the short run.

JEL Classifications: F41, F32, E52, E62

Keywords: Current Account, Exchange Rate, Risk Premium, Balance Sheet Effect

1. INTRODUCTION

The role of current account imbalance in determining macroeconomic crisis of emerging market economies has been a major concern in recent literature. The weaker macroeconomic fundamentals make an emerging market economy more vulnerable to a variety of internal and external shocks, which in turn, cause exchange rate depreciation with immediate worsening trade balance. The deterioration of current account balance in response to exchange rate depreciation can be attributed to the lagged response of exports and imports, that is, the J-curve phenomenon. This J-curve phenomenon has been observed for the developing countries like Egypt, Nigeria, Bangladesh, Malaysia, and Pakistan.¹ The present paper develops a structuralist framework to explore the dynamics of current account balance and exchange rate in conjunction with endogenous risk premium. This type of framework can be applied to a large class of emerging market economies which are subject to macroeconomic imbalance and the associated increase in

Moumita Basu <3.moumita@gmail.com> is Assistant Professor, Department of Economics, Krishnagar Government College, Krishnagar, Nadia, West Bengal, India. Ranjanendra Narayan Nag <rnag12@gmail.com> is Associate Professor, St. Xavier's College (Autonomous), Kolkata-700016, India. Bhaskar Goswami <mailbhaskar08@gmail.com> is Associate Professor, Department of Economics, University of Burdwan, India.

¹See Qurat-ul-Ain and Tufail (2013).

risk premium acting as trigger of macroeconomic crisis. Aronovich (1999) investigated the behaviour of the country risk premium for Argentina, Brazil and Mexico from June 1997 to September 1998. He found that the level of country risk premium is determined by several factors namely the US dollar bond market structure; restrictions on the acquisition of emerging market bonds imposed by the developed nations regulators; the credit risk measured by the notion of implied risk-neutral probability default; the different ways agents react to country risk due to asymmetric and imperfect information. The empirical evidence shows that the decrease in the country credit rating causes higher international borrowing cost.² In this context, we should mention that most of the developing countries are net debtors in the world market and are subject to underdeveloped financial market as compared to the developed countries, measured by credit worthiness indicators or by the quality of prudential regulation and supervision of financial institution.³ These lead to a risk premium in credit contracts of developing countries. Krugman (1999, 2003) emphasised the role of firm's balance sheets in explaining financial crisis in South-East Asian Countries. Sikdar (2014) explained why Indian asset market is volatile and identified factors which lead to dampen the optimism of the foreign investors about India that resulted in massive capital flight in accordance with high risk premium.

This paper is an attempt to address the issue of simultaneous response of current account, exchange rate and the output level to macroeconomic policies namely monetary policy and fiscal policy. In particular, we use the J-curve phenomenon for the analysis of dynamics of current account balance. The modelling of current account balance in this paper is significantly different from what is done in the existing literature.

This paper offers a structuralist treatment of different complications of adjustment in exchange rate.⁴ Current account responds to exchange rate with a time lag and evolves continuously. However, the exchange rate adjusts instantaneously in the asset markets. The adjustment in exchange rate produces a balance sheet effect on investment, current account, and risk premium as well. The balance sheet effect of exchange rate depreciation (Krugman, 2003) arises due to currency mismatch if the firms borrow in foreign currency but earn revenue in domestic currency. If there is a significant portion of foreign currency denominated debts, which is a true feature for the emerging market economies, then depreciation in the exchange rate leads to a rise in real debt burdens. Furthermore, such a rise in real debt occurs with a disproportionate increase in ability to pay and hence, produces adverse balance sheet effect on investment demand. Moreover, the paper includes an important structuralist feature of a developing country, namely wage price rigidity which in turn leads to Keynesian unemployment. Here, we address the issue of effective demand problem to explore dynamic interaction between current account balance and exchange rate.

The rest of the paper has been organised as follows. A brief review of literature is provided in Section 2. The theoretical model is explained and thoroughly discussed in Section 3. The comparative static exercises are examined in the following section. Section 5 offers summation of results along with concluding remarks.

²This was the case of lower rated Latin American countries' bonds. See Aronovich (1999).

³See Priewe (2008).

⁴The exchange rate is an asset price and its determination can be explained in a stock-flow consistent effective demand model.

2. RELATED LITERATURE

Scholars have explored and debated the interaction between exchange rate depreciation and current account imbalance since 1940. At that time, the elasticity pessimism view suggested that actual trade elasticities were too low to satisfy the Marshall-Lerner condition for trade balance improvement due to exchange rate depreciation. The trade balance immediately worsens following exchange rate depreciation as export and import quantities remain unchanged in the short run. However, overtime export demand begins to increase and import demand tends to fall. Hence, trade balance gradually improves in the long run as the elasticities rise. The tendency of the elasticities to rise overtime leads to phenomenon of the J-curve. There is voluminous empirical evidence regarding the phenomenon of the J-curve effect in developing countries. For instance, Lal and Lowinger (2002) confirmed existence of J-curve effect in their study of East Asian countries. Petrović and Gligorić (2010) have shown that exchange rate depreciation in Serbia improves trade balance in the long run, while giving rise to a J-curve effect in the short run. Qurat-ul-Ain and Tufail (2013) have explored the existence of Marshall-Lerner condition and J-Curve phenomenon for each D-8 countries. Gebeyehu (2014) also empirically verified existence of J-curve effect in Ethiopia. Using Vector Error Correction Model, Prakash and Maiti (2014) have found the evidence of J-curve phenomenon on goods trade, but not on services over the period 1975 to 2012.

Inclusion of implications of macroeconomic policies for the interaction between current account and exchange rate has been central to the recent macroeconomic studies. For instance, Bjornlan (2008) has analysed the transmission mechanisms of monetary policy in a small open economy like Norway through structural VARs, with special emphasis on the interdependence between monetary policy and exchange rate movements. Ferrero, Gertler, and Svensson (2009), and Makanza and Dunne (2015) examined implications of monetary policy and fiscal policy for current account dynamics. Abbas, et al. (2010) has investigated the relationship between fiscal policy and the current account, drawing on a large sample of advanced, emerging and low-income economies and using a variety of statistical methods including panel regressions (an analysis of large fiscal policy and current account changes) and panel vector autoregressions. Danmola, et al. (2013) have empirically examined the impact of monetary policy on the components of current account for the period 1970–2010 in Nigeria and their study confirmed a long-run relationship between monetary policy and components of current account under consideration. Buyangerel and Kim (2013) have analysed the effects of various macroeconomic shocks namely monetary policy shocks, price level shocks, output shocks and exchange rate shocks on trade balance and exchange rate in South Korea by using a structural vector error correction (SVEC) model. Prasad (2015) has assessed the role of capital controls, sterilised interventions and an exchange rate peg in explaining China's current account surplus.

This brief literature review shows that there is a dearth of theoretical research on current account imbalance and exchange rate dynamics in presence of endogenous risk premium and output adjustment.⁵ Blanchard (2005) and Dornbusch (1980) emphasised the role of net debt accumulation in explaining interaction between current account

⁵A notable exception is work by Blanchard, et al. (2005). This study, they explained how the US current deficits since the mid-1990s through the working out of the J- curve.

dynamics and the exchange rate. Our theoretical study, however, emphasises the role of endogenous risk premium in explaining interaction between current account imbalance and exchange rate dynamics in response to selective macroeconomic policies alongside output adjustment.

3. THE MODEL

The major building blocks of this model have been explained as follows:

Output is demand determined. The economy has an investment demand which is inversely related to both real interest rate and real exchange rate. The inclusion of real exchange rate as a determinant of investment demand is based on the balance sheet effect of exchange rate depreciation. There are three assets: domestic bonds, foreign bonds and money. Foreign bonds and domestic bonds are imperfect substitutes due to presence of risk premium. The disposable income is considered to be a determinant of consumption expenditure. The consumer price index depends only on exchange rate at fixed price. In this paper, the current account adjustment is expressed as the outcome of J-curve phenomenon, that is, lagged response of the trade balance to the exchange rate depreciation. On the other hand, the interest rate differential and risk premium together play an important role in determining the exchange rate dynamics under rational expectation.

3.1. The Commodity Market

In the commodity market, aggregate demand (AD) consists of expenditure on consumption (C) and investment (I), government expenditure (G) and trade balance (B).⁶

$$AD = C(Y - tY) + I\left(r, \frac{eP^*}{P}\right) + G + B \quad \dots \quad \dots \quad \dots \quad \dots \quad (1)$$

Consumption expenditure can be expressed as a positive function of disposable income ($Y - tY$). Investment demand is an inverse function of both real exchange rate $\left(\frac{eP^*}{P}\right)$ ⁷ and interest rate (r). Government expenditure (G) is considered to be exogenous. The aggregate output (Y) is demand determined and hence, the commodity market equilibrium is:

$$\begin{aligned} Y &= AD \\ \Rightarrow Y &= C(Y - tY) + I\left(r, \frac{eP^*}{P}\right) + G + B \quad \dots \quad \dots \quad \dots \quad \dots \quad (2) \end{aligned}$$

From Equation (2) we can determine the equilibrium output as:

$$Y^* = Y\left(r, \frac{eP^*}{P}, t, G, B\right) \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (3)$$

Let us explain partial effect of each variable on the output level. An increase in interest rate reduces investment demand leading to decrease in output level, that is, $Y_1 = \frac{\partial Y}{\partial r} < 0$. An increase in real exchange rate raises real value of external debt burden which in turn generates a negative balance sheet effect on investment demand and hence, output

⁶In this model we consider trade balance as equivalent to the current account balance.

⁷Here, P and P^* are domestic and foreign prices respectively. Both prices are assumed to be fixed.

falls to restore the commodity market equilibrium, that is, $Y_2 = \frac{\partial Y}{\partial \left(\frac{eP^*}{P}\right)} < 0$. An increase in tax rate reduces disposable income and hence, consumption expenditure. This leads to decrease in output level and hence, we get $Y_3 = \frac{\partial Y}{\partial t} < 0$. An increase in government expenditure raises output level such that $Y_4 = \frac{\partial Y}{\partial G} > 0$. A trade balance improvement causes an increase in output level, that is, $Y_5 = \frac{\partial Y}{\partial B} > 0$.

3.2. The Money Market

The money market equilibrium is:

$$\frac{M}{Q(eP^*, P)} = L(r, Y) \text{ [where, } L_1 = \frac{\partial L}{\partial r} < 0 \text{ and } L_2 = \frac{\partial L}{\partial Y} > 0] \quad \dots \quad \dots \quad (4)$$

The left-hand side of Equation (4) represents supply of real money balances. It can be expressed as the money supply deflated by consumer price index (Q) which in turn positively depends on domestic price (eP^*) of importables and price (P) of domestic goods. The right-hand side of Equation (4) shows the demand for money which is inversely related to the interest rate and positively related to output level.

From money market equilibrium we determine the equilibrium interest rate as:

$$r = r(e, Y, M, \dots) \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (5)$$

Let us explain partial effect of each variable on interest rate. An increase in exchange rate raises consumer price index which in turn reduces real money balance leading to excess demand in the money market. Hence, interest rate increases to restore money market equilibrium such that $r_1 = \frac{\partial r}{\partial e} > 0$. An increase in output level leads to higher demand for money and hence, higher interest rate, that is, $r_2 = \frac{\partial r}{\partial Y} > 0$. An increase in money supply leads to excess supply in the money market which in turn entails an increase in interest rate and hence, $r_3 = \frac{\partial r}{\partial M} > 0$.

3.3. Current Account Adjustment

The current account adjustment is given by

$$\dot{B} = \varphi(\alpha e - B) \text{ [where, } \varphi \text{ and } \alpha > 0] \quad \dots \quad \dots \quad \dots \quad \dots \quad (6)$$

\dot{B} denotes overtime adjustment in trade balance, φ being the speed of adjustment and α is the parameter and

$$\alpha = \text{Elasticity of export demand} + \text{absolute value of elasticity of import demand} - 1.$$

We assume that Marshall-Lerner condition is satisfied and hence, $\alpha > 0$. Empirical evidences suggest that both exports and imports are subject to lagged response to exchange rate depreciation.⁸ The evidences have found that only about 50 percent of the full quantity adjustment takes place in the first years; 90 percent occurs in the first

⁸Here, we consider that both domestic price and world price are given for this small open economy. Hence, any change in nominal exchange rate is equivalent to change in real exchange rate.

five years. There are several reasons⁹ why the response of export and import quantities to exchange rate depreciation is greater in the long run than in the short run. In Equation (6) the adjustment in current account arises due to lagged response of trade balance (B) to the exchange rate (e) depreciation.

3.4. Interest Rate Parity Condition with Endogenous Risk Premium

One of the major departures of our model from Mundell-Fleming framework is the assumption of imperfect substitutability of domestic bonds and foreign bonds. The return on domestic bonds is domestic interest rate (r) while the return on foreign bonds is the sum of foreign interest rate (r^*), expected change in exchange rate¹⁰ $\left(\frac{\dot{e}}{e}\right)$ and endogenous risk premium (ρ). Now, arbitrage between domestic bonds and foreign bonds therefore implies:

$$r = r^* + \frac{\dot{e}}{e} + \rho(D, B) \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (7)$$

Equation (7) shows the interest rate parity condition. The endogenous risk premium¹¹ (ρ) is positively related to the government budget deficit (D) and inversely related to the trade balance (B) and it is given by:

$$\rho = \rho(D, B) \text{ [where, } D = G - tY] \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (8)$$

An increase in government expenditure raises government budget deficit (D) leading to increase in risk premium and hence, $\rho_1 = \frac{\partial \rho}{\partial D} > 0$. On the other hand, the budget deficit gets reduced with an increase in the output level or tax rate which in turn raises the tax revenue of government. This leads to decrease in risk premium. An increase in trade balance reduces risk premium, that is, $\rho_2 = \frac{\partial \rho}{\partial B} < 0$.

Based on the interest rate parity condition we can express overtime exchange rate adjustment as the gap between interest differential and risk premium, represented by the Equation (7')

$$\frac{\dot{e}}{e} = r - r^* - \rho(D, B) \quad (7')$$

Equation (7') illustrates that an increase in domestic interest rate raises return on domestic bonds and hence, to restore the interest rate parity condition exchange rate will increase overtime, that is, $\dot{e} > 0$. On the other hand, an increase in risk premium makes foreign bonds more attractive. Consequently, restoration of interest rate parity condition leads to overtime fall in exchange rate such that $\dot{e} < 0$.

⁹First, there is a lag due to imperfect dissemination of information, during which importers recognise that relative prices have changed. Secondly, there exists a lag in deciding to place a new import order. Thirdly, after a new import order has been placed, there may be production and delivery lags before it is filled. Fourthly, producers sometimes relocate their factories to the country where costs are lower because of exchange rate advantage, regardless of whether it is the home country of the producer or the country where the goods are sold. This leads to the longest delivery process.

¹⁰Given perfect foresight, expected depreciation is equal to actual depreciation.

¹¹We acknowledge the idea of endogenous risk premium to Sikdar (2014), who used it in an extended Mundell-Fleming model.

3.5. Dynamic Adjustment and Stability

On the basis of Equation (6), the current account adjustment can be expressed as a function of trade balance and exchange rate as shown in the Equation (9)

$$\dot{B} = f(B, e) \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (9)$$

Let us interpret partial effect of each variable on \dot{B} . It follows from Equation (6) that $f_1 = \frac{\partial \dot{B}}{\partial B} = -\varphi < 0$. On the other hand, an increase in exchange rate entails overtime improvement in current account balance¹² such that $f_2 = \frac{\partial \dot{B}}{\partial e} = \varphi\alpha > 0$.

From above interpretation it is clear that the exchange rate depreciation causes a gradual improvement in trade balance so as to keep the current account balance at its steady state level, that is $\dot{B} = 0$, in the long run. This relationship is depicted by the positively sloped $\dot{B} = 0$ schedule in Figure (1). Any point to the left (right) of the $\dot{B} = 0$ locus represents trade deficit (surplus) with $\dot{B} > 0$ ($\dot{B} < 0$).

Based on Equation (7'), the exchange rate dynamics can also be written as a function of both trade balance and exchange rate as represented by Equation (10).

$$\dot{e} = g(B, e) \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (10)$$

Let us again explain the partial effect of each variable on exchange rate dynamics. From the Equation (8) it is clear that a trade balance improvement leads to decrease in risk premium. Moreover, the increase in trade balance raises output level as obtained from the commodity market equilibrium in Equation (3). This higher output causes higher equilibrium interest rate as explained by Equation (5), that is, $r_2 > 0$. The increase in interest rate and fall in risk premium together explain the overtime increase in exchange rate and hence, $g_1 = \frac{\partial \dot{e}}{\partial B} = r_2 Y_5 - \rho_2 > 0$. An increase in exchange rate raises consumer price index leading to decrease in real money balance and hence, there generates excess demand in the money market. As a result, the domestic interest rate will increase, that is, $r_1 > 0$. The higher interest rate has positive influence on exchange rate dynamics. On the other hand, the balance sheet effect of exchange rate depreciation causes lower investment demand which in turn reduces output level as obtained from Equation (3), that is, $Y_2 < 0$. The lower output leads to higher budget deficit and hence, higher risk premium ($\rho_1 > 0$) which negatively influences exchange rate dynamics. Therefore, it is clear that exchange rate depreciation has ambiguous effect on the dynamic adjustment in exchange rate. However, saddle path stability of the stationary state requires that exchange rate must increase overtime in response to exchange rate depreciation such that $g_2 = \frac{\partial \dot{e}}{\partial e} = r_1 + t\rho_1 Y_2 > 0$, that is, the increase in interest rate must dominate the increase in risk premium.¹³

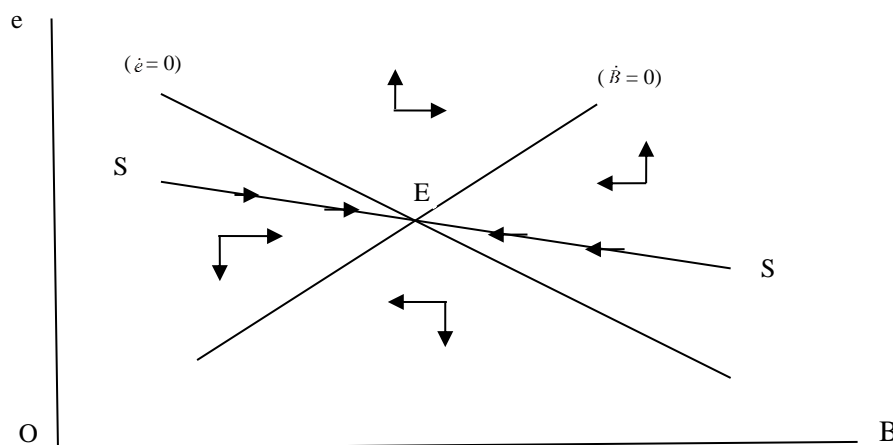
On the basis of above explanation, we can deduce that an increase in trade balance necessitates exchange rate appreciation so as to maintain the exchange rate at its steady state level, that is, $\dot{e} = 0$. It is represented by the downward sloping $\dot{e} = 0$ schedule in Figure (1).

¹²This is the essence of J-curve phenomenon.

¹³It implies that $r_1 > |t\rho_1 Y_2|$.

Any point that lies above (below) the $\dot{e} = 0$ locus, exchange rate would depreciate (appreciate) and $\dot{e} > 0$ ($\dot{e} < 0$)¹⁴ respectively.

Fig. 1. Phase Diagram Showing Saddle Path Stability



In Figure (1), the steady state equilibrium is represented by the point E at which $\dot{B} = 0$ and $\dot{e} = 0$.

Now, we concentrate on illustration of stable saddle path because it gives economically meaningful result. Assuming perfect foresight, unique convergent saddle path requires characteristic roots of one positive and one negative sign so that the determinant $|J| = \begin{vmatrix} f_1 & f_2 \\ g_1 & g_2 \end{vmatrix}$ is negative.

The saddle path SS is downward sloping and it is flatter than the $\dot{e} = 0$ schedule. For a specific trade balance level, the trajectory of the economy is from a particular value of exchange rate on the saddle path towards the stationary equilibrium, E. Here, exchange rate is a jump variable while trade balance is a slow moving variable.

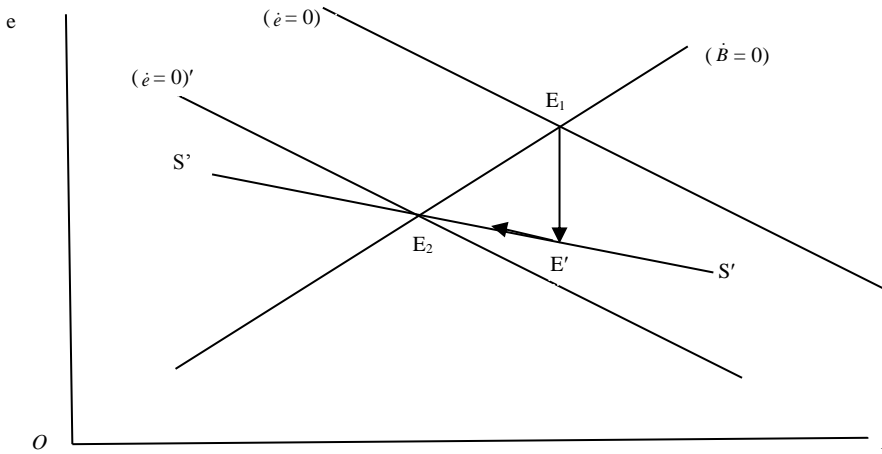
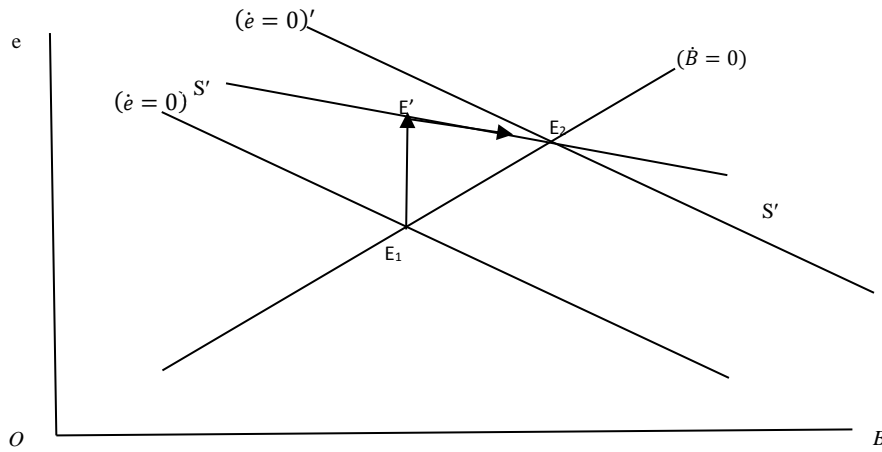
4. COMPARATIVE STATICS

This section analyses implications of fiscal policy and monetary policy for trade balance, exchange rate, and output.

4.1. Expansionary Fiscal Policy

Let us first consider an expansionary fiscal policy. The adjustment mechanism is as follows: the fiscal expansion leads to higher output level, higher demand for money and hence, higher interest rate which in turn has a positive effect on \dot{e} . On the other hand, the increase in government expenditure raises budget deficit which in turn causes an increase in risk premium. Clearly, it has negative influence on \dot{e} . Therefore, the expansionary fiscal policy has ambiguous effect on the exchange rate dynamics. If $\dot{e} > 0$, the $\dot{e} = 0$ schedule shifts downward as shown in Figure (2a). The opposite case will appear in Figure (2b).

¹⁴See appendix for derivations of the loci and saddle path stability.

Fig. 2a. Effect of an Expansionary Fiscal Policy with Exchange Rate Appreciation**Fig. 2b. Effect of an Expansionary Fiscal Policy with Exchange Rate Depreciation**

According to Figure (2a), in the short run, the movement from the point E_1 to the point E' on the new saddle path $S'S'$ is due to exchange rate appreciation. As exchange rate appreciates, we immediately get improvement in current account balance, that is, $\dot{B} < 0$ at the point E' since exports and import quantities would not change instantaneously. This is due to the presence of the pre-existing contracts whereby both the exporters and importers have to honour their prior trade commitments even though the exchange rate has appreciated. Thereafter, the subsequent trading activity reflects the new competitive situation with concomitant worsening trade balance in the long run as the economy gradually moves from the point E' to the point E_2 . The deterioration of trade balance moderates the initial decrease in exchange rate and hence in the short run, exchange rate overshoots. Corresponding to the new equilibrium point E_2 , the current account is balanced ($\dot{B} = 0$) and interest rate parity condition is maintained with steady state level of exchange rate ($\dot{e} = 0$). The expansionary fiscal policy along with exchange

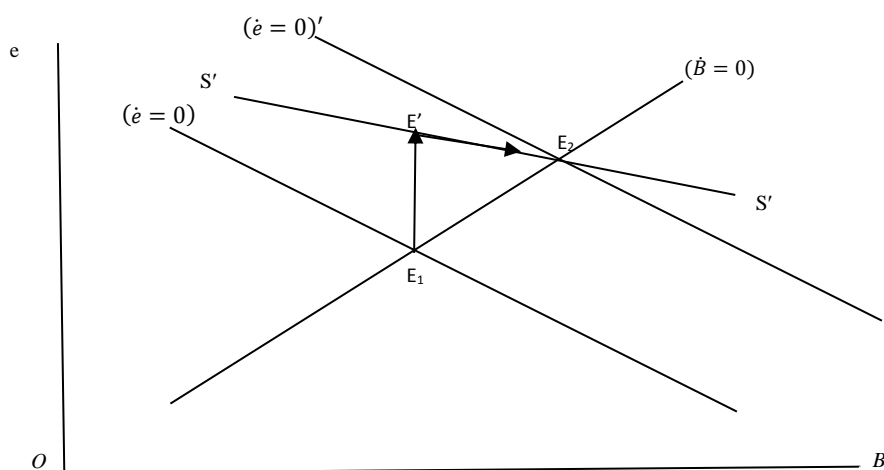
rate appreciation¹⁵ dominates the crowding out effect of higher interest rate leading to a short run increase in the output level and level of employment. However, in the long run, output as well as level of employment may fall if the effect of trade balance deterioration outweighs the balance sheet effect of exchange rate appreciation on investment demand.

In contrast, Figure (2b) shows that the exchange rate depreciates in the short run and the economy immediately jumps from the point E_1 to the point E' on the new saddle path $S'S'$ which shifts upward. The exchange rate depreciation immediately leads to trade balance deterioration, that is, $\dot{B} > 0$ at the point E' . In response to this exchange rate depreciation, overtime exports tend to increase and imports tend to fall leading to gradual improvement in trade balance so that in the long run the economy moves from the point E' to the point E_2 . The improvement in trade balance partially offsets the initial increase in exchange rate and hence in the short run, the exchange rate overshoots. The new steady state point E_2 represents balanced trade situation ($\dot{B} = 0$) and higher exchange rate compared to the old equilibrium point E_1 . In the short run, expansionary fiscal policy entails an increase in output level and level of employment since initial effect of fiscal expansion dominates the secondary crowding out effect of exchange rate depreciation along with higher interest rate. In the long run, the trade balance improvement reinforces the initial effect of fiscal expansion leading to increase in output level and level of employment.

4.2. Expansionary Monetary Policy

As money supply goes up, there is a fall in the domestic rate of interest, that is, the return on domestic bonds falls. This induces investors to switch their portfolio to foreign bonds leading to capital outflow. The decrease in domestic interest rate causes $\dot{e} < 0$. Consequently, the exchange rate depreciates leading to upward shift of the $\dot{e} = 0$ schedule as represented in Figure (3). However, the trade balance does not improve instantaneously following this depreciation. Nevertheless, the trade balance improves after a time lag along with an improvement in the current account.

Fig. 3. Effect of an Expansionary Monetary Policy



¹⁵The exchange rate appreciation generates a favourable balance sheet effect on investment demand.

Next, we turn to the dynamic adjustment process. An expansionary monetary policy leads to immediate jump of the exchange rate to point E' on the new stable path $S'S'$. Corresponding to the point E' the current account balance deteriorates in the short run, that is, $\dot{B} > 0$. The increase in exchange rate improves the trade balance overtime and the economy reaches the new stationary equilibrium point E_2 . The rise in trade balance offsets the initial increase in exchange rate and this explains why the exchange rate overshoots in the short run, as shown by point E' in Figure (3). The effect on output as well as on level of employment is ambiguous. If the balance sheet effect of exchange rate depreciation dominates other favourable effects of monetary expansion on output level, output contraction is inevitable in both the short and long run. Hence, level of employment falls. However, the long run contractionary effect is mitigated by the trade balance improvement. It is also to be noted that in the long run intensity of balance sheet effect diminishes since the steady state increase in exchange rate is less than the immediate effect.

5. CONCLUSION

The paper is an attempt to offer an explanation of dynamics of current account imbalance and that of the exchange rate in terms of a structuralist model in which output is demand determined and risk premium in the foreign exchange market is endogenous. While current account evolves continuously overtime, the exchange rate adjusts instantaneously in response to any shock which may be policy induced or otherwise. Though the empirical literature on the issue is copious, the analytical works do not always offer a very clear account of certain essential features of a developing country. In this context, one may refer to the paper by Blanchard, et al. (2005) that discusses the dynamics of current account in the context of US economy with a clear focus on the role of international investors. However, the paper abstracts from the problem of unemployment which may also arise due to effective demand constraint as well as the endogenous adjustment in risk premium in the foreign exchange market. Moreover, our paper examines the effect of exchange rate on aggregate demand not only in terms of the price effect on net exports but also the balance sheet effect of exchange rate depreciation on investment demand.

Among the major findings of our paper include the following:

Expansionary monetary policy results in depreciation of exchange rate. The immediate effect is the short run deterioration of trade balance. However, the steady state equilibrium corresponds to the improved trade balance. Moreover, the monetary expansion may reduce the output level and level of employment both in the short and long run if balance sheet effect of exchange rate depreciation offsets other favourable effects of monetary expansion on the level of output. The exchange rate overshoots in the short run. An expansionary fiscal policy has ambiguous effect on both exchange rate and trade balance. The fiscal expansion may cause output contraction and fall in level of employment if trade balance deteriorates in the long run.

We suggest a couple of extensions of the model. One possible extension is to recast the model in a dependent economy framework by introducing traded and non-traded goods. Moreover, issues of capital accumulation and debt dynamics can also be addressed.

APPENDIX

A.1: $\dot{B} = 0$ and $\dot{e} = 0$ locus:

The slope of $\dot{B} = 0$ locus can be derived as follows:

$$\dot{B} = f(B, e) \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (9)$$

$$a_{11} = \frac{\partial \dot{B}}{\partial B} = f_1 < 0$$

$$a_{12} = \frac{\partial \dot{B}}{\partial e} = f_2 > 0$$

Now slope of the $\dot{B} = 0$ curve is

$$\left. \frac{de}{dB} \right|_{(\dot{B}=0)} = -\frac{a_{11}}{a_{12}} > 0$$

The slope of $\dot{e} = 0$ curve is obtained from the following equation

$$\dot{e} = g(B, e) \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (10)$$

$$a_{21} = \frac{\partial \dot{e}}{\partial B} = g_1 > 0$$

$$a_{22} = \frac{\partial \dot{e}}{\partial e} = g_2 > 0$$

Now slope of the $\dot{e} = 0$ curve is

$$\left. \frac{de}{dB} \right|_{(\dot{e}=0)} = -\frac{a_{21}}{a_{22}} < 0$$

A.2: Saddle Path

The matrix of first partial derivatives for (9) and (10) is

$$J = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} f_1 & f_2 \\ g_1 & g_2 \end{bmatrix}$$

Since $f_1 < 0, f_2 > 0, g_1 > 0$ and $g_2 > 0$, $|J|$ is negative. As $|J|$ is negative, one of the characteristic roots must be negative for saddle point stability.

Time paths of B and e

$$B(t) = B^* + A_1 e^{\mu_1 t} + A_2 e^{\mu_2 t}$$

$$e(t) = e^* + \frac{a_{21}}{[\mu_1 - a_{22}]} A_1 e^{\mu_1 t} + \frac{a_{21}}{[\mu_2 - a_{22}]} A_2 e^{\mu_2 t}$$

with B^* and e^* being the equilibrium values of B and e; A_1 and A_2 are constants; and μ_1, μ_2 are two characteristic roots. Here we assume that $\mu_1 > 0$ and $\mu_2 < 0$.

Let $A_1 = 0$. Thus,

$$B(t) = B^* + A_2 e^{\mu_2 t}$$

$$e(t) = e^* + \frac{a_{21}}{[\mu_2 - a_{22}]} A_2 e^{\mu_2 t}$$

The equation that describes the saddle path is:

$$e(t) - e^* = \frac{a_{21}}{[\mu_2 - a_{22}]} [B(t) - B^*]$$

Slope of the Saddle Path

$$\left. \frac{de}{dB} \right|_{SS} = \frac{a_{21}}{[\mu_2 - a_{22}]} < 0 \quad [\text{since, } \mu_2 < 0]$$

Slope of the $\dot{e} = 0$ schedule

$$\left. \frac{de}{dB} \right|_{(\dot{e}=0)} = -\frac{a_{21}}{a_{22}}$$

Hence, the slope of saddle path is less than that of the ($\dot{e} = 0$) schedule.

A.2 Comparative Static Results

The steady-state effects of monetary policy and fiscal policy can be obtained from the following equations:

$$\dot{B} = f(B, e) = 0 \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (9')$$

$$\dot{e} = g(B, e) = 0 \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (10')$$

A.2.1. Expansionary Fiscal Policy:

The Steady State Effects:

$$\dot{B} = f(B, e) = 0 \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (9')$$

$$\dot{e} = g(B, e, M, G) = 0 \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (10')$$

Differentiating Equation (9') and (10') with respect to G and setting $\dot{B} = 0$ and $\dot{e} = 0$ respectively and arranging them in matrix form we get,

$$\begin{bmatrix} f_1 & f_2 \\ g_1 & g_2 \end{bmatrix} \begin{bmatrix} \frac{dB}{dG} \\ \frac{de}{dG} \end{bmatrix} = \begin{bmatrix} 0 \\ -g_G \end{bmatrix}$$

Applying Cramer's Rule we get,

$$\frac{dB}{dG} = \frac{f_2 g_G}{|J|} > 0 \quad \text{if } g_G = \frac{\partial \dot{e}}{\partial G} < 0$$

..... < 0, otherwise

$$\frac{de}{dG} = \frac{f_2 g_G}{|J|} > 0 \quad \text{if } g_G = \frac{\partial \dot{e}}{\partial G} < 0$$

..... < 0, otherwise

A.2.2. Expansionary Monetary Policy:

The Steady State Effects:

$$\dot{B} = f(B, e) = 0 \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (9')$$

$$\dot{e} = g(B, e, M, G) = 0 \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (10')$$

Differentiating Equation (9') and (10') with respect to M and setting $\dot{B} = 0$ and $\dot{e} = 0$ respectively and arranging them in matrix form we get,

$$\begin{bmatrix} f_1 & f_2 \\ g_1 & g_2 \end{bmatrix} \begin{bmatrix} \frac{dB}{dM} \\ \frac{de}{dM} \end{bmatrix} = \begin{bmatrix} 0 \\ -g_M \end{bmatrix}$$

Applying Cramer's Rule we get,

$$\frac{dB}{dM} = \frac{f_2 g_M}{|J|} > 0, \quad \text{since } g_M = \frac{\partial e}{\partial M} < 0]$$

$$\frac{de}{dM} = -\frac{f_1 g_M}{|J|} > 0 \quad \text{since } g_M = \frac{\partial e}{\partial M} < 0]$$

REFERENCES

- Ali, Abbas M. S., Bouhga-Hagbe, J., Fatas, A., Mauro, P., & Velloso, R. C. (2010). Fiscal policy and the current account. (IMF Working Paper, WP/10/21).
- Aronovich, S. (1999). Country risk premium: Theoretical determinants and empirical evidence for Latin American countries. *Rio de Janeiro*, 53(4), 463–498.
- Bjornlan, C. H. (2008). Monetary policy and exchange rate interactions in a small open economy. *The Scandinavian Journal of Economics*, 110(1), 197–221. DOI: 10.1111/j.1467-9442.2008.00532.x.
- Blanchard, O., Giavazzi, F., & Sa, F. (2005). International investors, the U.S. current account, and the Dollar. *Brookings Papers on Economic Activity*, 1, 1–49.
- Buyangerel, B. & Kim, W. J. (2013). The effects of macroeconomics shocks on exchange rate and trade balances in Korea. *Korea and the World Economy*, 14(1), 91–119.
- Caves, R. E., Frankel, J. A., & R. W. Jones (2006). *World trade and payments: An introduction*. Pearson Education, Inc. and Dorling Kindersley Publishing Inc.
- Danmola, R. Akonji, Olateju, & Olubunkola, A. (2013). The impact of monetary policy on current account balance in Nigeria. *IOSR Journal of Humanities and Social Science (IOSR-JHSS)* e-ISSN: 2279-0837, p-ISSN: 2279-0845. Volume 7, Issue 3 (Jan. - Feb. 2013), PP. 67–72.

- Ferrero, A., Gertler, M., & Svensson, L. E. O. (2009). Current account dynamics and monetary policy. *International Dimensions of Monetary Policy*, Galí and Gertler. 2009.
- Gebeyehu, B. A. (2014). Exchange rate and trade balance; J-Curve effect in Ethiopia. *Journal of Economics and Sustainable Development*, 5(24). ISSN 2222-1700 (Paper) ISSN 2222-2855 (Online)
- Krugman, P. (2003). Crises: The next generation? In Helpman, E. and Sadka, E. (eds.) *Economic policy in the international economy: Essays in honour of Assaf Raxin*. Cambridge University Press.
- Lal, K.A., and Lowinger, T. C. (2002). The J-Curve: Evidence from East Asia. *Journal of Economic Integration* 17(2), 397—15.
- Makanza, S. C., & Paul Dunne, J. (2015). Current account dynamics and monetary policy transmission in South Africa. *School of Economics Macroeconomic Discussion Paper Series*, School of Economics, University of Cape Town, July, 2015.
- Makanza, S. C. & Paul Dunne, J. (2015). Fiscal consolidation, fiscal policy transmission, and current account dynamics in South Africa. (ERSA Working Paper 565, December, 2015).
- Priewe, J. (2008). Capital account management or Laissez-faire of capital flows in developing countries. In Arestis, P., Paula, L. de, de Paula, & Fernando, Luiz (eds.) *Financial liberalisation and economic performance in emerging countries*. Palgrave Macmillan, Springer.
- Petrović, P., and Gligorić, M. (2010). Exchange rate and trade balance: J-curve effect. *PANOECONOMICUS*, 2010, 1, pp. 23-41. DOI: 10.2298/PAN1001023P.
- Prakash, K., and Maiti, D. (2014). The J-curve disparity in the goods and services sector in the pacific region: Evidence from devaluation in Fiji. NZAE 2014 Conference, at Auckland, New Zealand.
- Prasad, N. (2015). The role of macroeconomic policy in explaining China's current account surplus. https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=ACE2015&paper_id=275.
- Qurat-ul-Ain, S. and Tufail, S. (2013). The effect of oil price shocks on the dynamic relationship between current account and exchange rate: Evidence from D-8 countries. *The Pakistan Development Review*, 52(4).
- Sikdar, S (2014). Mundell-Fleming with stock market and endogenous risk premium. In Acharyya, R. & Marjit, S. (ed.). *Trade, globalisation and development*. Essays in Honour of Kalyan K. Sanyal: Springer.

Size and Age as Determinants of Employment Growth among Manufacturing Firms in Pakistan

FARRUKH IQBAL and AADIL NAKHODA

Size-based industrial policy (support for small firms) has long been provided by the government in Pakistan while age-based policy (support for young firms) has become prominent in recent years. Both policies are typically justified by reference to positive effects on labour absorption. Despite their popularity among policymakers, however, the empirical basis for such policies has not been adequately analysed at the national level. We address this issue using data from a large, multisector, random sample of manufacturing firms and find empirical support for size-based policies but not for age-based ones. We also find that size-based policies appear most relevant for firms with less than fifty workers.

JEL Classification: Firm Performance: Size, Diversification and Scope (L25);
Industrialisation. Manufacturing and Service Industries.
Choice of Technology (O14)

Keywords: Firm Size, Firm Age, Employment Growth, Manufacturing; Pakistan,
Industrial Policy

1. INTRODUCTION

In many countries, one important objective of industrial policy is to promote small firms. Fiscal, financial and infrastructure subsidies are offered to small firms and this is usually justified by one or both of two arguments: that such firms deserve help on a fairness basis (equity motivation) or because helping them has positive externalities for national employment, productivity or output growth (efficiency motivation). Such policies often carry a direct financial (or fiscal) cost and may also carry an indirect and longer-term economic cost associated with the distortions introduced by discriminatory incentives. It is important, therefore, to have a good empirical sense of the link between firm size and the relevant policy objective in order to justify size-based industrial policies in a specific country context.

In the case of low income countries, a policy of positive discrimination in favor of small firms is sometimes justified by reference to their superior allocative efficiency with respect to labour. Smaller firms employ more labour per unit capital and this is

Farrukh Iqbal <fiqbalus@gmail.com> is Director, Center for Social Policy Research, Karachi. Aadil Nakhoda is Assistant Professor of Economics, Institute of Business Administration, Karachi.

Authors' Note: This paper was presented at a webinar organised by the Pakistan Institute of Development Economics (PIDE) in September 2020. We are grateful to the participants of the webinar for useful comments and suggestions.

considered to be more efficient because low income countries tend to be relatively labour-abundant. Strictly speaking, better alignment with national factor endowments is not sufficient to show superior efficiency. For this, smaller firms have to be more efficient with respect to the use of all factors, in other words, to have higher total factor productivity. This assumption is not uniformly supported by empirical evidence. Indeed, there is much evidence that suggests the opposite, that larger firms have higher levels and growth of total factor productivity as, for example, in the cases of Japan (Urata and Kawai, 2002) and Taiwan (Aw, 2002).

Another rationale rests on the ability of smaller firms to exploit dynamic economies. They are agile and adapt faster to changing market demands. Since they typically operate in competitive sectors with low barriers to entry, they have to innovate more in order to survive. Some evidence for this view is provided by Acs and Audretsch (1990) and Audretsch (2002) who show that the patenting rate for small firms in the US is typically higher than that for larger firms measured on a per employee basis.

However, the above justifications focus on the relationship between size and productivity and this need not translate into an equivalent link between size and employment growth. For the latter, a better rationale may be found in the stylised lifecycle of the typical firm. Most firms start out small, often based on the initiative of a motivated entrepreneur with limited funds. Over time, such firms become larger as owners get more experience, as they reinvest net earnings and as they benefit from economies of scale. This process continues until a stage where diseconomies of scale set in. Expanding the firm further risks dilution of management attention and control. Most owner-managed firms plateau at this level. In this conceptual model, size drives employment growth dynamics. However, a plausible story could also be told with age as the driver of employment growth. In this alternative model, a firm grows fast when it is young, based on the effort, ideas and risk-taking typical of young owners and managers. At some point, growth begins to slow as the owners/managers become more risk-averse with age and avoid increasing value at risk.¹ This too would create an employment growth profile featuring rapid growth at younger ages and flatter growth over time.²

Which version is correct? The one in which job growth is driven by size or that in which it is driven by age? The answer is important for designing and implementing industrial policy. At present, policymaking in most countries is dominated by popular acceptance of the version in which size drives job growth, although some age-based policies have also been introduced in some countries.³ Ultimately, disentangling size and age effects on employment growth requires careful empirical analysis. Such analyses have been conducted for developed countries (see Evans, 1987 and Haltiwanger, et al.

¹On the tendency for people to become more risk averse with age, see Dohmen, et al. (2017).

²A similar profile can be derived from models (e.g., Jovanovic, 1982) which define output as a function of management efficiency and link the latter to age. The basic idea is that owner/managers learn about their firm's relative efficiency over time; firms that prove to be less efficient exit while the more efficient remain in business and grow larger. Most of this discovery of efficiency takes place when the firm is young and is reflected in higher firm size volatility. As firms age, there is less discovery and less volatility. One implication of such a learning process is that younger firms show more dynamism (growth in size) than older firms.

³It appears that the attention of Pakistani policymakers is also shifting in this direction. Recent editions of the Pakistan Economic Survey (for 2018 and 2019) devote more paragraphs to youth-based policies and development schemes than to size-based ones. These policies and schemes typically relate to employment and skills.

2013 for the US) as well as for developing countries (see Bigsten, et al. 2007 and Ayyagiri, et al. 2014). For Pakistan, only one study has tackled this issue (Wadho, et al. 2019) but it applies to the textiles and apparel sector only. The main contribution of the present paper is an empirical analysis of size and age effects across all manufacturing sectors for Pakistan using a national random sample of establishments.

Pakistan offers several incentives based on firm size. The State Bank of Pakistan has programs that affect the supply of credit to small firms through commercial banks or development finance organisations. These generally take the form of lending quotas, credit guarantee and refinance schemes and interest rate subsidies, but often go beyond.⁴ The Small and Medium Enterprise Development Authority (SMEDA) provides free or subsidised business development services to small firms.⁵

An inverse relationship between firm size and employment growth has long been noted for both developed and developing countries. For the US this relationship was highlighted in a paper by Birch (1981) and has since been periodically reconfirmed, including most recently in Neumark, et al. (2011). A similar inverse link is reported for low income countries (for recent references, see Aterido, et al. 2011 and Ayyagiri, et al. 2014). Pakistani studies in this domain tend to be mostly concerned with the contributions of the entire category of small and medium firms to total employment. This approach, however, confuses mass with dynamism. Almost everywhere, small firms are more numerous than large firms and account for the bulk of employment. To show that smaller firms are more dynamic requires firm-level data and analysis.⁶

The joint assessment of size and age also goes back decades, with early theoretical contributions by Boyanovic (1982) and empirical tests for US data by Evans (1987) showing firm employment growth to be inversely linked to both size and age. More recently, Haltiwanger, et al. (2013) reported that the inverse link with size disappears in the US data once firm age is controlled for. Young firms were found to have the fastest rates of job creation, thus highlighting the role of startups in the employment picture. Among studies for developing countries, Ayyagiri, et al. (2014) find that both size and age matter: smaller and younger firms have higher rates of job creation than larger and more mature firms. In the case of Pakistan, Wadho, et al. (2019) also find both size and age to be inversely related to employment growth but, as noted earlier, their study is confined to only one manufacturing sub-sector (textiles and apparel).⁷

⁴The relevant page on the State Bank's website (<http://www.sbp.org.pk/70/sup-14.asp>) notes that its Policy for Promotion of SME Financing (2017) covers areas like "regulatory relaxations, financing targets, provision of refinance and risk coverage facilities, promotion of value chain financing and program based lending, adoption of technology, awareness creation and capacity building of bankers as well as SMEs, handholding of SMEs and facilitative taxation regime for SMEs."

⁵The SMEDA website (<https://smeda.org/>) notes the following among its objectives for small and medium enterprises: facilitation of business development services, helping small firms get financing, providing training, assisting in attainment of international certification, coordinating external donor assistance and conducting sector studies.

⁶Sur, et al. (2014) report results from a national sample of rural non-farm enterprises in Pakistan in which they examine the role of investment climate variables and include size. They report a negative coefficient for size but their dependent variable is output (or value-added) rather than employment growth.

⁷Waheed (2017) includes age (but not size) as a control variable in an empirical study of the effect of innovation on employment growth in Pakistan.

In attempting to clarify the roles of size and age in firm dynamics in Pakistan, we are specifically interested in the following questions: Is there a robust relationship between size and employment growth in the presence of firm age? Is there a robust relationship between firm age and employment growth in the presence of size? How do these relationships unfold across different size and age groups?

To generate results that can provide guidance for national level policies, we need a database that has at least the following four characteristics: it is national in scope, collected as a random sample, focused on the establishment as the unit of reporting, and contains data on employment growth at the establishment level.⁸ For Pakistan, the latest available database with these four characteristics is the World Bank Enterprise Survey conducted during 2013-15 (WBESP, 2013). This survey was administered to owners or top managers of a representative sample of formal (registered) private non-agricultural firms from all over Pakistan.⁹ This is the database we use, focusing only on manufacturing sector firms.

The remainder of this paper is organised as follows. Section 2 lays out the empirical strategy we follow. Section 3 discusses the results of the empirical analysis. Section 4 contains some concluding remarks on the implications of our results for the design of industrial policy in Pakistan.

2. EMPIRICAL STRATEGY

Defining the dependent variable. The WBESP provides information on current employment at the time of the survey (L1) and employment two years ago (L2). Accordingly, we can calculate the percentage change in employment between L1 and L2, divide this by 2 to get an annual rate, and use this as our dependent variable. Using logarithms, this may be calculated as below:

$$L_{ijr} = (\log(L1) - \log(L2))100/2 \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (1)$$

Adjusting for fixed effects. Regression results based on cross-section data are subject to the influence of many unobserved variables. One way to offset the effect of some unobserved variables is to use the “fixed effects” of some known exogenous proxies. In our present paper, we use industry/sector and region as these proxies. Since industrial policies (tax breaks, financial subsidies, infrastructure access etc.) often vary by sectors and province, including these proxies can help account for such variation.

Estimating Strategy. The OLS form of our basic model is as follows:

$$L_{ijr} = \beta_1 (size_{ijr}) + \beta_2 (age_{ijr}) + \delta_j + \theta_r + \varepsilon_{ijr} \quad \dots \quad \dots \quad (2)$$

where L_{ijr} is the employment growth rate for the i th firm, in industry j belonging to region r . The number of employees in the base year (L2) is denoted by $size_{ijr}$ and the age of the firm at the time of the survey is denoted by age_{ijr} . ε_{ijr} is the error term, δ_j denotes industry fixed effects and θ_r region fixed effects.

⁸While we use the terms interchangeably in the paper, our analysis is conducted at the establishment and not at the firm level. More than one (location-defined) establishment may be part of the same firm.

⁹The survey covered 1247 firms stratified by industry, establishment size and geographic region. After removing unreliable and inaccurate responses (as determined by the enumerator) and selecting only manufacturing firms, the dataset was left with 971 usable observations. A description and discussion of the survey methodology is accessible at: <https://www.enterprisesurveys.org/en/methodology>

We estimate this model for three size groups and three age groups. Recall that the stylised model provides a basis for dividing the growth of a firm over time into at least two and possibly three phases: an early phase, where due to small size or young age, it grows rapidly; a middle phase where the growth rate moderates and a latter phase where employment may even decline. One testable implication of this model is that the employment growth rate among larger firms will be lower than that among smaller firms and the employment growth rate among older firms will be lower than that among younger firms. In other words, there is an inverse relationship between size (or age) and employment growth.

Theory provides no guidance to delimiting the appropriate size and age groupings. The convention for many developing countries is to group firms by size as follows: Small, for firms having less than 20 workers; Medium, for firms having between 20 and 99 workers; and Large, for firms having 100 or more workers. This convention is used by the WBESP in reporting its results and by several recent studies (see Aterido, et al. 2011 and Ayyagiri, et al. 2014). We adopt it for this paper as well. As far as age is concerned, we define the following three groups: Youngest, for firms 10 years or less in age; Young, for firms between 11 and 20 years in age; and Mature, for firms more than 20 years in age.¹⁰

Robustness checks through control variables. We then subject the basic model to robustness checks through the inclusion of multiple control variables. The literature suggests that, beyond size and age, employment growth is typically linked to two sets of factors, one external to the firm and the other internal to it. The external set refers to the environment in which firms operate and the internal to owner/manager characteristics and preferences. The following aspects of the investment climate of a country are usually found to be important among external factors: infrastructure (especially transportation and power); regulatory burden (especially as expressed through tax administration); and access to finance. While the relevant literature for Pakistan is sparse, the importance of most of the above factors is confirmed by information in the WBESP survey that identifies what firms self-report as the most important constraints they face.¹¹

Accordingly, we select the following variables to control for the investment climate in Pakistan: availability of electricity; quality of tax administration; and access to finance. The availability of electricity is measured by whether or not the firm had a generator or shared one. The burden of tax administration is measured by whether or not the establishment received a visit from a tax official during the survey year. The assumption is that such a visit would have involved the payment/collection of a bribe. Access to finance is measured by whether or not the firm had an active loan or a line of credit.

We also select control variables that relate to choices made by owners/managers regarding participating in exporting, training for employees, obtaining international certification, and generating innovative products. All of these are reported as binary

¹⁰Ayyagiri, et al. (2014) identify their youngest group, Startups, as being between 0 and 5 years of age. Such firms form only 3 percent (28 observations) of our sample. Results based on such a small fraction of the sample would not have been credible; hence we use a larger age group (up to 10 years) to generate a comparison of age/employment growth slopes over time.

¹¹For example, 75 percent of the respondents in the sample identified the availability of electricity as a leading constraint and 34 percent reported tax administration processes as an impediment.

variables. Participating in exporting measures whether or not the firm exports any amount of its output. Training indicates whether or not the firm offers in-service training to its staff. International certification measures whether or not the firm had obtained ISO 9000 and/or related certifications. Innovation measures whether or not the firm self-reports producing and marketing a new product.

The OLS form of the fuller model including multiple control variables is as follows:

$$L_{ijr} = \beta_1 (size_{ijr}) + \beta_2 (age_{ijr}) + \beta_3 Controlvariables_{ijr} + \delta_j + \theta_r + \varepsilon_{ijr} \dots (3)$$

Summary Statistics. Summary statistics (see Table 1) reveal quite a lot of variation in the size of firms, ranging from a minimum of 1 and a maximum of 15000, though the median firm is small at 20 employees. There is also a lot of variation in firm age which ranges from a minimum of 2 to a maximum of 77. The median firm is a relatively mature 20 years.¹² Among firm characteristics of interest, we note that as many as 35 percent report having international certification which is more than twice the percentage that report engaging in exporting. This suggests that many firms obtain international certification for advantages or benefits that apply in the domestic market. Twenty-two percent also report providing training opportunities to their workers. Among business climate variables of interest, we note that 57 percent report having been visited by a tax official in the survey year and 53 percent report owning or sharing a generator. This is consistent with the general sense among respondents that these factors are important constraints to doing business in Pakistan. We also note that only 9 percent report having a new loan or a line of credit which suggests limited access to finance. Finally, we note that as many as 29 percent report having introduced a new product, reflecting attempts at innovation in a competitive environment.

Table 1

Summary Statistics

	Mean	Median	Std. Dev.	Min	Max
Employment Growth	5.47	0.00	18.62	-101	129.51
Number of Employees in Base Year	166.86	20.00	770.94	2	15000
Age of Firm	23.03	20.00	13.77	2	77.00
International Certification	0.35	0.00	0.48	0	1.00
Formal Training	0.22	0.00	0.41	0	1.00
Exporter	0.16	0.00	0.37	0	1.00
New Loan or Line of Credit	0.09	0.00	0.29	0	1.00
Owned or Shared a Generator	0.53	1.00	0.50	0	1.00
Visit by Tax Officials	0.57	1.00	0.49	0	1.00
Introduced a New Product	0.29	0.00	0.46	0	1.00

¹²Almost three-quarters of the sample consists of sole proprietors with the owner very likely being the top or key manager as well. The median experience of top managers is 15 years which is three quarters of the median age of firms and suggests low turnover in this category.

Correlations

Table 2 provides a subset of the correlation matrix focusing on the three variables of most concern to this study, namely, employment growth, firm size and firm age. No strong patterns stand out.

Table 2
Selected Correlations

	Employment Growth	Number of Employees in Base Year	Age of firm
Employment Growth	1.00		
Number of Employees in Base Year	-0.02	1.00	
Age of Firm	-0.08	0.13	1.00
International Certification	0.13	0.22	0.26
Formal Training	0.08	0.22	0.14
Exporter	0.00	0.16	0.12
New Loan or Line of Credit	-0.04	-0.03	-0.04
Owned or Shared a Generator	-0.08	0.16	0.22
Visit by tax officials	-0.16	0.08	0.01
Introduced a New Product	-0.13	0.03	0.17

3. EMPIRICAL RESULTS AND RELATED DISCUSSION

The Basic Model

We estimate the model as follows. First, we run a model in which only size and age feature, along with industry and region fixed effects. In this version, both size and age are entered as continuous variables but in logarithmic form to minimise the effects of outliers. The results are shown in column 1 of Table 3. Second, we disaggregate the sample by size and age. The smallest size category and the youngest age category are excluded in the regression. The results are shown in column 2 of Table 3.

Table 3
Impact of Size and Age on Employment Growth

Dep. Var: Employment Growth	(1)	(2)
Number of Employees in Base Year(ln)	-2.15*** (0.44)	
Age of firm (ln)	-0.52 (1.15)	
Medium-sized Firms		-9.54*** (1.86)
Large Firms		-8.08*** (1.68)
Younger Firms		-3.32 (2.42)
Mature Firms		-1.95 (2.40)
Constant	12.86*** (3.72)	11.47*** (2.26)
Observations	804	804
R-squared	0.10	0.13

Robust standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1.

Includes industry and region fixed effects.

The results show that firm size is inversely related to employment growth but firm age is not. As firms grow larger, employment growth gets smaller but the same does not appear to happen with age. This result holds even when the sample is disaggregated. Employment growth for medium and large firms is smaller than it is for small firms (the excluded group). A similar pattern does not apply across the age groups. This result is different from that of Ayyagiri, et al. (2014) where both size and age are shown to matter for firms in developing countries. It is also different from Wadho, et al. (2019) where both size and age are reported to be inversely related to employment growth among textile and apparel firms in Pakistan. It is, however, consistent with Waheed (2017) in which an insignificant result is reported for age using a national multi-sector sample.¹³

Robustness Checks

We check for robustness by adding a set of control variables to the basic model as shown in Table 4. The results suggest that our basic model is robust. The sign and

Table 4

Effect of Size and Age in Presence of Multiple Control Variables

Dep. Var: Employment Growth	(1)	(2)
Number of Employees in Base Year(ln)	-3.40*** (0.70)	
Age of Firm (ln)	-2.28 (1.43)	
Medium-sized Firms		-10.17*** (2.07)
Large Firms		-12.12*** (2.77)
Younger Firms		-4.01 (2.82)
Mature Firms		-4.15 (2.95)
International Certification	11.33*** (2.18)	9.96*** (2.08)
Formal Training	7.20*** (2.24)	6.40*** (2.17)
Exporter	1.24 (2.39)	1.38 (2.43)
New Loan or Line of Credit	-0.12 (2.77)	-0.06 (2.72)
Owned or Shared a Generator	-0.95 (1.68)	-1.11 (1.68)
Visit by Tax Officials	-8.22*** (1.74)	-8.21*** (1.74)
Introduced a New Product	-4.29** (1.81)	-3.38* (1.81)
Constant	26.49*** (5.04)	18.33*** (3.02)
Observations	675	675
R-squared	0.20	0.21

Robust standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1.

Includes industry and region fixed effects.

¹³We have also estimated a version of the basic model in which the non-linearity of size and age effects is tested through the use of the squares of the size and age variables (in their logarithmic forms). The results are similar to those reported in Table 3: size has a significant quadratic relationship with employment growth but age does not have a significant relationship.

significance of the size and age variables do not change with the addition of a large set of control variables. Employment growth remains inversely linked with size but not with age. This is true whether size and age are taken in the aggregate or separated into groups.

Brief Comments on Control Variable Results

While the control variables are not of principal interest to this paper, a few remarks may be in order to show how they compare with results found in other studies. With regard to the variables that reflect firm characteristics, we note that getting certification and providing training to workers are positively and significantly related to employment growth. This is consistent with our priors. We are unaware of any studies for Pakistan that show such results. We also note that innovation is negatively associated with firm dynamism. This is contrary to the results reported by Waheed (2017) who finds a positive link between product innovation and employment growth among Pakistani manufacturing firms, though only for those in low-tech sectors. The difference in our results may be due to differences in our estimating strategies. Waheed (2017) models the innovation-dynamism link as an endogenous one and estimates it via a two stage procedure. It also differs from Wadho, et al. (2019) in which various measures of innovation are found to be positively linked to employment growth among Pakistani textile and apparel firms.¹⁴

As far as the business environment variables are concerned, we note that the quality of tax administration turns out to have a significant adverse effect: visits by tax officials are associated with lower rates of employment growth. This is also reported for Pakistan by Abbas, et al. (2020). Similar, though more nuanced, results have been reported for corruption variables by Aterido, et al. (2011) for developing countries as a whole. Infrastructure, as measured by the availability of a generator, and access to finance, as measured by having a loan or a line of credit, do not show up as significant. On infrastructure, we note that Aterido, et al. (2011) find a negative link between the incidence of power outages and employment growth for small, medium and large firms but a positive link for micro firms. They interpret this as indicating that micro firms use less energy in their activities and are not sensitive to the availability of power. We also note that Ahmed and Hamid (2011) report a positive link between access to finance and employment growth in Pakistan. However, they consider this link to be endogenous and estimate it using a two stage procedure, which we have not done. For further guidance on this particular issue, we would urge the reader to consult that study.

4. FINDINGS AND POLICY IMPLICATIONS

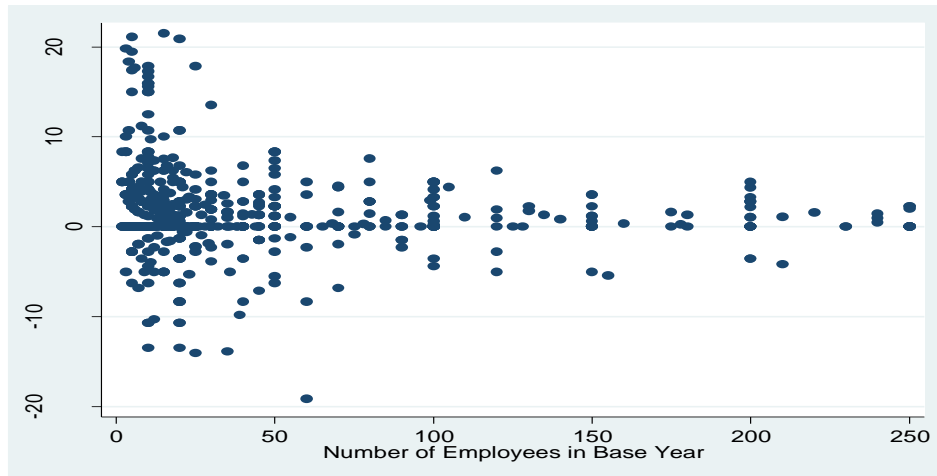
We find two main results. First, firm size is a statistically robust determinant of employment growth in Pakistan. As firm size increases, the rate of employment growth declines. Second, firm age does not have a statistically significant impact on employment growth. In this section, we explore some additional aspects of these findings.

¹⁴The negative coefficient on product innovation is consistent with a monopolistic competition view. If product innovation allows firms to enjoy some pricing power in local markets, they could respond by raising prices and lowering output. The lower output would be consistent with slower employment growth. We have also tried a dummy variable reflecting process innovation instead of product innovation in the full regression: the variable turns out to be statistically insignificant.

Exploring the Link between Firm Size and Employment Growth

We first look at the scatter plot displayed by the data when we put firm size on the horizontal axis and employment growth on the vertical axis (see Figure 1 below.)

Fig. 1. Scatter Plot of Firm Size against Employment Growth



The scatter plot recreates in graphical form the main result for size that we have established through statistical analysis earlier in this paper, namely, a negative relationship between firm size and employment growth. In addition, it shows that employment dynamism is concentrated at the smaller end of the size distribution where firms have less than 50 employees or so. After that, the slope is mostly flat.¹⁵

The scatter plot also shows considerable churning among smaller firms who not only generate jobs faster but also destroy jobs faster. Small new firms face difficult odds and many of them suffer steep job losses even while others report rapid job gains. Some, indeed, may not survive the challenges of establishing a new business. This pattern is widely observed among firms in both developed and developing countries. It sets up a challenge for policy makers who, when faced with an application for benefits under some government program, must assess whether the applicant firm will create more jobs than it destroys over some period of time. This requires additional data and research on the link between firm characteristics, contextual considerations and job creation outcomes. In other words, though it is reasonable to start with it, the criterion of size alone does not provide a sufficient basis for a confident decision to provide, say, financial support.

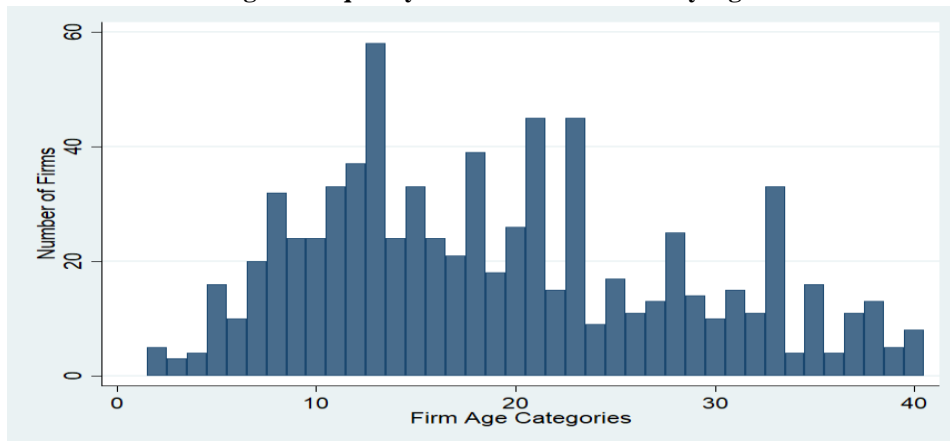
Why Does Firm Age not have an Impact on Employment Growth in Pakistan?

Firm age does not turn out to be significant in any of the variations of our model. This is surprising in view of the significance revealed in other studies noted for the US as well as for developing countries as a whole. We examine this matter further by looking at the distribution of firms by age in our sample.

¹⁵Figure 1 is based on 766 observations because we have excluded firms having more than 250 workers in order to be able to focus on the detail at the smaller-size end of the scatter plot.

Figure 2 shows that our sample contains very few very young firms. Only around 3 percent or so of the sample are firms at or under 5 years of age, a limit often used to define start-ups. In many countries, net job creation is highest among such start-ups. With relatively few start-ups in our sample, we should expect less employment dynamism on average. The relevant question then is why are there so few start-ups among manufacturing firms in Pakistan? The answer may lie in the costs associated with registering formally. Many firms avoid registering because they are afraid of attracting attention from the tax authorities and other government agencies. It may be only after they are older and well-established that they perceive the benefits of registration to outweigh the costs.¹⁶ This aspect of start-up dynamics in Pakistan needs to be further investigated through better data.

Fig. 2. Frequency Distribution of Firms by Age



Pakistan also appears to have a low entrepreneurship rate. Information on this is available through the Global Economic Monitor (GEM) database for Pakistan (2012). According to this, our total early stage entrepreneurship activity rate for opportunity-based (rather than need-based) entrepreneurship was just over 3 percent. This is only one third as much as the 9 percent average found in other low income countries in the GEM sample. Furthermore, only 3.4 percent of respondents reported being established as new business owners for up to three and half years. This compares unfavourably with an average survival rate of 13 percent in other low income countries in the GEM sample. These characteristics show entrepreneurship to be relatively weak in Pakistan. This is consistent with our finding of an insignificant contribution from young firms to overall employment dynamism: not enough such firms are being created by Pakistani entrepreneurs and/or not enough are surviving past their early years.¹⁷

¹⁶Almost one quarter of the firms in our sample report not having been registered formally when first established. This is consistent with our finding of a negative impact on firm growth arising from the burden of dealing with the tax authorities.

¹⁷A more recent survey on Pakistani entrepreneurship was conducted in 2019. This shows a total early stage entrepreneurship activity rate of 3.7 percent which places Pakistan 49th among the 50 countries that participated in 2019. The relevant country report for Pakistan has not been published but summary results are available in the global report (see Global Entrepreneurship Monitor, 2020).

One further observation is relevant. While we have not found age to be a significant determinant in the full sample, it may be significant for sub-samples focusing on specific sectors and types of firms. For example, Wadho, et al. (2019) report that smaller, younger and innovative firms exhibit much higher employment growth than the sample average in Pakistan's textiles and apparel sector.¹⁸ This is a useful finding that offers a path out of the policy dilemma faced when size alone is the benefits-granting criterion. If policy-makers had access to additional information such as the age, innovation status and sub-sector of applicant firms, they should be able to make better decisions.

REFERENCES

- Abbas, M., Faridi, M. Z. and Rahman, S. (2020). Assessing the factors affecting growth of small firms in Pakistan. *European Journal of Applied Business and Management*, 6(1), 77–89.
- Acs, Z. J. and Audretsch, D. B. (1990). *Innovation and small firms*. Cambridge: MIT Press.
- Ahmed, H. and Hamid, N. (2011). Financing Constraints: Determinants and Implications for Firm Growth in Pakistan. *Lahore Journal of Economics*, 16: (Special Issue), 317–346.
- Aterido, R., Hallward-Driemeier, M., and Pagés, C. (2011). Big constraints to small firms' growth? Business environment and employment growth across firms. *Economic Development and Cultural Change*, 59(3), 609–647.
- Audretsch, D. (2002). The dynamic role of small firms: Evidence from the US. *Small Business Economics*, 18(1-3), 13–40.
- Aw, B. Y. (2002). Productivity dynamics of small and medium enterprises in Taiwan. *Small Business Economics*, 18(1-3), 69–84.
- Ayyagiri, M., Demirguc-Kunt, A. and Maksimovic, V. (2014). Who creates jobs in developing countries? *Small Business Economics*, 43:1, 75–99.
- Bigsten, A. and Gebreeyesus, M. (2007). The small, the young and the productive: Determinants of manufacturing firm growth in Ethiopia. *Economic Development and Cultural Change*, 55(4), 813–840.
- Birch, D. (1981). Who creates jobs? *Public Interest*, 65, 3–14.
- Dohmen, T., Falk, A., Golsteyn, B., Huffman, D., and Sunde, J. (2017). Risk attitudes across the life course. *Economic Journal*, 127(605), F95–F116
- Evans, D. (1987). The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. *The Journal of Industrial Economics*, 35(4), 567–581.
- Global Entrepreneurship Monitor (2012). Pakistan Report. Accessed at: <https://www.gemconsortium.org/file/open?fileId=49071>
- Global Entrepreneurship Monitor (2020). Accessed at: <https://www.gemconsortium.org/file/open?fileId=50443>
- Haltiwanger, J., Jarmin, R. and Miranda, J. (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics*, 95(2), 347–361.
- Jovanovic, B. (1982). Selection and evolution of industry. *Econometrica*, 50, 649–70.

¹⁸The mean net employment creation in the combined category of small, young and innovative firms is reported as being more than 6 times higher than the sample average (Wadho, et al. 2019, Table 7).

- Neumark, D., Wall, B., and Zhang, J. (2011). Do small businesses create more jobs? New evidence for the United States from national establishment time series. *Review of Economics and Statistics*, 93, 16–29.
- Pakistan, Government of (2018-19). *Pakistan Economic Survey (2018-19)*. Accessible at: http://finance.gov.pk/survey/chapters_19/Economic_Survey_2018_19.pdf
- Sur, M., Zhiang, J. and Chen, K. (2014). The investment climate and enterprise performance in rural Pakistan: Implications for rural non-farm employment generation. *Journal of the Asia Pacific Economy*, 19, (3).
- Urata, S. and Kawai, H. (2002). Technological progress by small and medium enterprises in Japan, *Small Business Economics*, 18(1-3), 53–67.
- Wadho, W., Goedhuys, M., and Chaudhry, A. (2019). Young, innovative companies and employment creation: Evidence from the Pakistani textiles sector. *World Development*, 117, 139–152.
- Waheed, Abdul (2017). The employment effect of innovation: Evidence from Bangladesh and Pakistan. *The Pakistan Development Review*, 56(2), 105–126.
- WBESP (World Bank Enterprise Survey for Pakistan) (2013). Accessible at <http://www.enterprisesurveys.org/data/exploreeconomies/2013/pakistan>

Acemoglu, Johnson and Robinson’s Notion of Exogenous Imposition of Colonial Institutions onto Colonies— A Critique in the Light of Historical Evidence

MADEEHA GOHAR QURESHI, UNBREEN QAYYUM,
MUSLEH UD DIN, and EJAZ GHANI

This paper provides critique of Acemoglu, Johnson, and Robinson’s (2001, 2002) notion that rests on the hypothesis of exogenous imposition of colonial institutions onto their respective colonies based on conditions for their settlement. Our research brings forth the logical loopholes in Acemoglu, Johnson, and Robinson (AJR) by constructing arguments against the over-simplified assumption of exogenous imposition of colonial institutions in explaining the differences in development today. To prove our point, we build on two main arguments from history to show that some degree of endogeneity did indeed exist in colonial institutions that were imposed on the colonies. Our first argument revolves around the theme that how Atlantic slave trade evolved with colonialism and had meaningful technological and institutional consequences in the colonial metropolitan state. And these evolving conditions in coloniser’s mother country not only shaped incentives for mercantilist colonialism at one level and at the other became the base of institutional setup of progressive forms. In our second part of the argument, we demonstrate the role of native agency either in the form of local’s formal or informal pre-colonial institutions or in the form of their hold within the colonies, were all important in shaping what path colonisers eventually took for the institutional transfer. Based on these historical evidences, it is concluded that colonial institutions cannot be assumed as an exogenous transfer based on the notion of settlement as per AJR, rather it can be best described as an evolving fit between colonial and pre-colonial institutions.

Keywords: Acemoglu, Johnson, and Robinson’s (2001, 2002), Reversal of Fortune, Institutions, Colonialism, Long-term Development Differences

1. INTRODUCTION

Acemoglu, Johnson, and Robinson’s (AJR) research on the reversal of fortune has been counted as the seminal contribution in applied institutional economics.¹ According to their studies, the institutional course of action colonisers adopted in their colonies was based on two critical instruments namely settler mortality rate (Acemoglu, Johnson & Robinson, 2001)² and population density of indigenous people encountered by European

Madeeha Gohar Qureshi <madeeha.qureshi@pide.org.pk> is Research Economist, Pakistan Institute of Development Economics, Islamabad. Unbreen Qayyum <unbreenqayyum@pide.org.pk> is Staff Economist, Pakistan Institute of Development Economics, Islamabad. Musleh ud Din <musleh.ud@gmail.com> is Ex-Joint Director, Pakistan Institute of Development Economics, Islamabad. Ejaz Ghani <ejazg930@gmail.com> is Professor of Economics, Lahore School of Economics, Lahore.

¹Other important empirical studies on institutions and growth link that came about the same time AJR studies came out includes Hall and Jones (1999), Keefer and Knack (1997), Rodrik and Subramanian (2003).

²The empirical findings in Acemoglu, Johnson and Robinson (2001) provide evidence for above hypothesis and estimate robust negative and significant impact of settler mortality on the indicators of good institutional performance in 1900 and explain the phenomenon of reversal of development pattern post-colonisation through impact of colonisation on the institutional set up of the regions depending on the mortality they faced during colonisation.

colonist at the time of the conquest (Acemoglu, Johnson, & Robinson 2002).³ As per their hypothesis, areas where Europeans had the possibility of developing large scale settlement, they promoted European style inclusive institutional setup. In contrast, the colonies where they found little scope of placing their people, they established an extractive institutional mechanism. Hence, finding evidence in favour of the above hypothesis, AJR not only identified instruments that could explain the difference in institutional setup and long-run economic development of various colonies but also brought forward a beautiful blend of development economics and economic history.

Their work is unique not only in terms of findings but also in its innovative use of econometrics in establishing the impact of institutional paths taken by the colonial administration on post-colonial development prospects. Following their methodological breakthrough, many applied papers have tried to link post-colonial outcomes to its colonial origins, which include (Angeles, 2007; Bairoch, 1993; Banerjee and Iyer, 2005; Bertocchi, and Canova, 2002; Engerman & Sokoloff, 1997, 2005a, 2005b; Ferguson & Gupta, 2002; Grier, 1999; Iyer, 2010; Mohoney, 2010; Nunn, 2008a, 2008b and Reno, 1995).

In the context of the validity of AJR formulation, many studies have tested the robustness of these critical instruments. For instance, Alouy (2004a, 2004b) and Slyweter (2008) examined the robustness of settler mortality as the causal determinant of contemporary institutions after controlling for other historical and geographical correlates. Similarly, some other studies have tested the validity of AJR hypothesis by extending methodologies or by assessing the strength of data which includes McArthur & Sachs (2000), Auer (2009), Albouy (2012), Acemoglu, Johnson & Robinson (2012), Easterly & Levine (2016) and Chakraborty (2017).⁴ However, in our study, we have tried to augment the AJR line of research by going beyond econometrics justifications to an evaluation of these instruments in the light of historical evidence.

Hence, our work distinguish from previous comments on AJR that instead of targeting on the problems in settler mortality rate's measurements as others have done, our work advances the AJR argument that yes indeed the European colonial institutions have made difference in defining development path for their colonies but one should not ignore the question that what may have shaped the colonial institutions after 1500? In this context, we build an argument that how local people's choices within colonies and the evolution of institution⁵ of slavery had a meaningful role to play.

³Both these channels are related than it seems, as we know from work on historical conditions of development in Diamond (1999) that areas with sparse population neither had the capacity to develop diseases that could adversely impact possibilities of European settlement nor the strength to withstand European military might. While areas with denser native inhabitants had the ability to sustain more complex societies with internal strength to survive even in the face of colonialism and adverse disease environment for the European immigrants making settlements harder to come by (Bardhan, 2005, pp. 510–511).

⁴McArthur & Sachs (2000) test the robustness of AJR results to the change of sample both in size and geographical dispersion; Auer (2009) disentangles the partial effects of institutions and of endowments on income; Albouy (2012) criticise the authenticity of AJR's dataset and Acemoglu, Johnson & Robinson (2012) provides answers to criticism raised by Albouy (2012); Easterly and Levine (2016) extends the literature in line with AJR work by constructing a new variable—the European share of the population during colonisation and examining its link with development patterns today. Chakraborty (2017) not only points to measurement errors but test the AJR hypothesis of institutional primacy and geographical simultaneously.

⁵Given AJR use property right expropriation as institutional proxy; hence our notion of institutions will be the same. However, since extraction of colonisers was both in terms of materials and in case of Africa indulges in slave trade so we will define property rights expropriation in terms of both material extraction and slavery (primarily in case of Africa) as was categorised by AJR.

The key findings from our analysis are two-fold. First, we showed that how evolution in Atlantic slave trade with colonialism had meaningful technological and institutional consequences in the colonial metropolitan states defining their incentives for mercantilist and progressive colonialism differently. And secondly, imposed colonial institutions on the colonies can only be best described as an evolving fit between colonial and pre-colonial institutions with a defining role of native agency in sustainability of the final outcome. Hence, some degree of endogeneity did exist where conditions in colonies other than those related to settlement for the colonisers⁶ had indeed played an important role in defining the institutional path taken by colonial powers within their colonies.

Our findings have important relevance in understanding the process of liberal imperialism⁷ with foreign aid acting as a measure of influence on the developing countries (Hayter, 1971; Williams, 2000, 2008; Harvey, 2003; Itmi (2018); Williamson, 2018). Learning from Altaf (2011), Lum umba-Kasongo (2011) and Itmi (2018), one finds that not only foreign aid has become a means of imposing neo-imperialism on developing world but its continuation is in the interest of all stakeholders from aid donors to aid recipient countries and donor agencies like IMF and World Bank etc. For this purpose, the institutional weakness within the aid receiving countries serves both as source of aid inflow and also of its continuation (Easterly & Pfitze, 2008; Haque, 1999, 2018, 2020).⁸ This process has parallel to our findings. Just as in case of our analysis, pre-colonial institutions set-up of colonies defined the trajectory of colonial rule, in case of foreign aid the institutional weaknesses of the developing countries which often are rooted in their colonial past serve as the cause of aid inflow in its first stage and as source of further weakening of their institutions in its second phase. This creates a recurrent need to look out for aid and loans within developing countries undermining their future long-term growth prospects.

In this perspective, let us present empirical and historical evidence in favour of our hypothesis in Section 2 below. In light of this discussion, we will conclude the debate in final Section 3.

2. PLAUSIBLE LOOPHOLES / ANTITHESIS TO AJR EXPOSITION FROM HISTORICAL EVIDENCE

AJR, in their path-breaking contributions, not only have given us tools to integrate economic history and development economics but also have provided us with the most widely accepted explanation for development differences today. However, despite its remarkable contribution in explaining disparities in development patterns across the world using colonisation as a natural experiment, we showed how exogenous imposition of colonial institutions based on criterion of European settlement or non-settlement has undermined some possible sources of endogeneity for institutional transfers via

⁶Given we have tried to understand process of colonialism and institutional path within colonies in the backdrop of AJR framework, hence in our analysis wherever coloniser is used, it refers to only the era of European colonialism and all other forms of colonial experience such as the Romans, the Byzantines, the Moors, and the Ottomans, to name a few by default are being categorised as pre-colonial.

⁷Liberal Imperialism is being used in the sense of interventionist foreign policy by United States (Bishai, 2004; Mabee, 2004; Kiely, 2005; Ikenberry, 2006; Quinn and Cox, 2007).

⁸For example, Haque (1999) shows that reforms have failed in Pakistan because the intent was just to get money from donor to release immediate fiscal constraints. Further, this process of foreign aid and donor's assistance has led to creation of intellectual and institutional gaps as shown in Haque (2018a, 2018b).

colonialism. First of such source that we highlighted deals with how technological frontier for the colonial states evolved as prices of African slaves changed with colonisation defining the incentives for mercantilist and progressive colonialism differently. AJR did not take into account the role of local agency and pre-colonial institutions in shaping the form of colonialism they encountered. For more details we will divide the coming discussion into following two segments.

- 2.1. Evolution in Technological Frontier and Institutions of the European Colonial Powers and Role of Atlantic Slave Trade.
- 2.2. Argument of Local People as a Maker of their History.

2.1. Evolution in Technological Frontier and Institutions of the European Colonial Powers and Role of Atlantic Slave Trade

To establish endogeneity argument, we will first assess whether classification into settler and non-settler colonies as identified by AJR is the only valid demarcation to explain colonial institutional transfers, or there exist further sub-divisions within colonial experiences. Revealing one such division in section 2.1a, we build the argument of how such divisions are independent of AJR's criterion of coloniser's settlement or non-settlement for institutional transfer. We further discussed the institutional path that European Colonisers choose to administer their respective colonies was endogenous and highly dependent on how technological frontier for the colonial states evolved as prices of African slaves changed with colonisation and consequently their institutions, defining their stakes for mercantilist and progressive colonialism differently.

2.1a. *Notions of Colonialism Independent of Settler Criterion*

Diversity in colonial policies across various regions and the time dimension, due to varied political and economic institutional make of colonisers in their home countries need to be assessed within a rigorous historical account. In the above aspect, two distinct forms of colonialism pattern emerged i.e. mercantilist and progressive colonialism⁹ which

⁹Here we have just considered two divisions that is mercantilist and progressive colonialism but a thorough review of literature also points towards further finer divisions such as those shown in analyses of colonialism by Marx and Bill Warren (progressive but not liberal) (Warren, 1973; Marx, 1976; Amin, 2018), by Paul Baran and Lenin (neither progressive nor liberal) (Lenin, 1917; Baran, 1957) and (progressive and liberal) (Melleuish, 2001; Sang, 2009, 2018). However, ignoring these finer divisions does not undermine validity of our analysis since all we are trying to show that colonialism can be categorised in divisions other than the AJR notion. And that such divisions are motivated by reasons other than AJR notion of coloniser's settlement or non-settlement. However, both these critiques will stand valid even if we go beyond mere mercantilism or progressive divide to include some other subdivision etc. However, for simplicity of the argument, we keep to only two categories of mercantilism or progressive forms in our analysis. Examples of liberal and progressive colonies include British Hong Kong and British Australia (Melleuish, 2001; Sang, 2009, 2018). However, if we go deeper into philosophy of liberalism applied to these regions, then even in across these two example, we can find some import distinctions in its application. In Australia, the idea of liberalism is actually adopted in its British settlers who wanted more state intervention to guide policies towards certain end. In contrast, in Hong Kong, liberal policies were promoted within British colonial rule to promote a free market mechanism by design to protect it from other ideologies present into its neighborhood (Melleuish, 2001). Example of progressive colonialism Includes British North American Colonies and British Canada (Bourinot, 1900; Wrong & Langton, 1914; Frankema, 2010). Example of progressive but not liberal colonialism include the case of South Korea as colony of Japan—though this example is Non-European and does not apply to our context but for understanding of the process, it is important case study (Kohli, 2004; Hundt & Bleiker, 2007). Most prominent example of neither progressive nor liberal is case of British India (Wilson, 2016; Tomlinson, 1981; Tharoor, 2017; Lenin, 1917).

were not related to AJR notion of settlement criterion for inclusive institutional transfer and non-settlement benchmark for extractive institutional implantation.

The first form of colonialism, termed as mercantilism¹⁰ in literature, is based on short-term gains and restricted economic activity under colonialism (Coleman, 1957). In this form of political economy, the colonial focus was the maximisation of extraction at the lowest cost possible. And the prime emphasis was on only immediate consumption goals and maintenance of colonial supremacy with no economic rationality for investment to increase the productive base of the colonies for long-term sustainable progress.¹¹ Hence, the administration within this form of colonialism implemented state regulations that benefited the coloniser's merchants, their aligned beneficiaries, and support groups both within colonies and in their home country (Ekelund & Tollison, 1980a, 1980b).

To sum up, the mercantilist colonialism patronised economic stratification within its colonial subjects officially. These practices had resulted in their weak institutional base and eventually led to negative post-colonial prospects for regions that had a history of being subjected to such colonial plunder. On the other hand, colonies under progressive colonialism operated in an atmosphere where free market forces could kick in their true spirit at domestic and international level.

Colonial powers within liberal agendas made sure that commercial interests should not be politically influenced within their subjects. State involvement was restricted to the extent that the incentives for entrepreneurial investments and technical up-gradation through property rights protection were maintained. The only hierarchy that resulted within such colonies was market forces and not through official policy of divide and rule as practiced within mercantilist orientation.

A case-oriented approach to the analysis of historical evolution processes reveals that few colonial governments were found to adopt the mercantilist style of ruling i.e. British India¹² and Spanish American colonies (Wilson, 2016; Tomlinson, 1981 and Tharoor, 2017; Engerman & Sokoloff, 1997, 2005b) while others who were liberal in their approach indulged in progressive capitalism since its very inception for example consider the case of British North America (Frankema, 2010), British Canada (Bourinot, 1900; Wrong & Langton, 1914) and British Australia (Melleuish, 2001). Hence, contrary to AJR's hypothesis, we found variation in line of progressive or mercantilist approach to ruling in the colonial history both within colonies and the colonial powers.

¹⁰Since some historians use concept of mercantilism as a representation of all sorts of non-laissez-faire economic practices in European history from fifteenth through nineteenth centuries, hence to clarify such confusion it should be noted that terms mercantilist and progressive phase of colonialism being used in present discussion should be recast as classical mercantilism and enlightened mercantilism respectively when making current context comparable to other works employing above definition of mercantilism being applied for whole of the colonial history.

¹¹The imperative towards technological improvements of an economy and training of the workforce was avoided within the colonies under the mercantilist rule, given it would have absorbed the returns of the colonial extraction (Mohaney, 2010).

¹²Analysis by Wilson (2016), Tomlinson (1981) and Tharoor (2017) shows that different facets of British colonialism and its extraction.

However, for the sake of clarity, one should note that the AJR notion of extractive and inclusive settlement should not be confused with the mercantilist and progressive phases of colonialism, respectively. The exploitative colonial rule within AJR analysis was linked only to the unfeasibility of settlement. In contrast, the vision for extraction differed across mercantilist and liberal colonialism was not in light of settlement criterion but more so in the face of the technological frontier of the time and institutional setup within coloniser home country. Hence there could be both phenomena of more or less settler presence across colonies experiencing a mercantilist or liberal form of colonialism (Mohaney, 2010).

Further, given that technological frontier faced by colonists evolved with colonisation, hence taking this feedback impact from colonial encouragement of extractive practices on themselves needs to be understood. Given colonisation not only affected the conditions in colonies but as these conditions evolved so did the technological and institutional structures within the coloniser's mother country and with its objective of colonialism and its form.¹³ For example, Britain that led the transition towards the capitalism through initiating industrial revolution could only do that as the balance of power changed between monarchies, commercial, and business groups post Atlantic slave trade¹⁴ (Acemoglu, Johnson & Robinson, 2005, Inikori, 2002). Hence what sort of institutions a colonial state had imposed on its colonies, did in some sense was determined by changing conditions and contributions of colonies into how the colonial metropolitan country evolved itself institutionally and hence this process was not exogenous as AJR assumed. AJR claim that eight European nations¹⁵ that had participated in the process of colonialism imposed a common institutional framework in their respective colonies depending on whether they had established massive settlements or not is easily challengeable and unstable.

2.1b. *Evolution in Technological Frontier and Institutions within Colonial Metropolitan States with Colonisation and Role of Atlantic Slave Trade*

The strategies for colonisation depended not only on the extent of settler mortality and native population density alone but also on internal dynamics and the stage of development at which the colonial powers had been at the time of its indulgence into colonial projects. In the initial phase of the colonial takeover near 1500, the technological frontier had not evolved for the European counties to the extent that oppressive institutions of mercantilists could change. The practice of labour-intensive techniques for

¹³The extent of capitalism within the societies of colonial powers also had an essential role in defining their colonial trajectories. Moreover, these institutional differences among the European colonial powers had a consequential impact on the route such forces took under their colonial mission. Countries that could not evolve as effective capitalist institutions for themselves cannot be the cause of leading their colonies towards such growth-promoting end. The possibility of positive institutional transfers in the sense of free-market norms and *laissez-faire* economic policy with colonialism could happen only where such practices had emerged in the mother country.

¹⁴Rough estimates of profits from British slave trade during eighteenth century comes up to be 50 million pounds, from West Indian sugar plantations to be between 200 to 300 million pounds and from Indian subcontinent between 500 million to one billion pounds (Crow & Thorpe, 1988, p. 16; Digby, 1969, p. 33). These profits did play an important role in creating enabling environment to kick start industrial revolution in Britain which then spread all over Europe subsequently (Inikori, 2002).

¹⁵European nation that indulged in overseas territorial control through colonialism at some point between fifteenth and mid-twentieth centuries include Belgium, Britain, France, Germany, Italy, Netherlands, Portugal, and Spain.

the plantation of cash crops gave a big push to forced labour (Engerman & Sokoloff, 1997). Slavery was considered as the most productive and feasible option given the technology employed for production was not safe and health-friendly (Fenoaltea, 1984; Fogel & Engerman, 1974). With the established practice of human pawning in pre-colonial Africa, the abundance of African slaves at a cheap cost (Austin, 2005) and technological improvements in the shipping industry made the Atlantic slave trade feasible.

From the mid-eighteenth century onwards, massive outflow of African slaves, along with substantial demographic loss from ongoing tribal wars within Africa, pushed the price of these slaves up¹⁶ (Stavrianos, 1981). Such resulting unmet slave demand, along with their price rise, created incentives for their replacement with mechanised production (O'Rourke, and Williamson, 1999; Acemoglu, Johnson, & Robinson, 2005). Further, this shift away from labour extensive techniques towards a transition for a capital oriented technological innovation forced a change in the class structure in some of the coloniser's metropolitan states e.g. in Britain, Netherland and Holland to name a few (O'Rourke, and Williamson, 1999, Inikori, 2002). These changes had multiple impacts. For example in Britain, At some level the increasing tendencies for capitalism post such changes initiated a search for territorial expansions and new markets that set a base for a process of extractive colonial rule in India for Britain as has been stressed by Marx and Lenin (Kohn, & Reddy, 2017). And at other level, with changing class structures by strengthening of merchants groups in Britain post Atlantic slave trade, and a shift of labour from rural to urban areas (industrialisation phase) led to weakening of feudalism in Britain hence initiating a process inclusive institutional reforms (Acemoglu, Johnson, & Robinson, 2005; Rogowski & Macae, 2004). Further, how the Atlantic slave trade transformed the institutional path of the colonial countries was not uniform and depended on to whom revenue of this trade went to. In Britain and Netherland, since the beneficiaries were the merchant groups and in Spain and France, revenues went directly to the crown, hence the institutional evolution in these areas differed greatly (Acemoglu, Johnson, & Robinson, 2005). Such feedback impact over long-term evolution cannot be accommodated within AJR quantification and can be best understood under comparative historical analysis.

In light of the above discussion, few historical facts should be given their due merit in defining the inter-linkages for development prospects across the world in the face of changing conditions for colonialism over time. Firstly, proprietorship in people through institutions of slavery in Africa responded to evolving technological constraints such as the 'cash-crop revolution' from the initial colonial timeframe (Austin, 2005, pp. 236–249, 512–515). Secondly, in context of how the contribution of this oppressed workforce to mining industries for gold, diamonds, coal, minerals, etc. or to the

¹⁶The demographic loss due to colonial indulgence into the slave trade and death toll that can be attributed to resulting tribal warfare or in the process of shipment of slaves onto western hemisphere during 1600-1900 time period has been approximated to be 12 million and 36 million (Stavrianos, 1981, p. 109). In percentage terms, there was a drop of ten percentage points from 18 percent to 8 percent in African share in the world population from 1650 to 1850, a large portion of which can explain directly or indirectly through the phenomenon of Atlantic slave trade (Stavrianos, 1981).

development of vast irrigation networks¹⁷ had provided a stimulus to the industrial revolution globally within the European colonisers.¹⁸ Such a complete picture of how extractive institutions in one region had contributed positively towards development in other areas has not been given due significance in the AJR assumption of exogenous institutional transfers from coloniser to colonies.

2.2. Argument of Local People as Makers of their History

So far, we have focused on the institutional diversity of the coloniser and its impact on the colonies; however, a more problematic puzzle that AJR story is unable to shed light on is that why do we find variation in institutional adoption even across colonies under same colonial administration? More precisely, why does one find regions that were both ruled under colonisers belonging to similar school of thought¹⁹ show diversity in their institutional path and hence post-colonial outcomes? Such historical variety within colonies under the same colonial institutional exposure needs to be assessed in the light of emerging evidence using an interdisciplinary approach involving sociologists, economists, and historians' point of views taken together. The dynamics of indigenous populations in choosing their historical path and the strength of pre-colonial institutions as defining factors for such divergent colonial trajectories will be presented in more detail as below:

2.2a. *Evolving Fit of Pre-conquest Institutions and Colonial Institutions*

AJR discourse builds on the premise of heavy settlement as a key element for positive institutional transfers from the metropolitan state to its colonial constituencies. However, the true process of colonialism and institutional impact involves more dynamism than has been accommodated in AJR methodology. So far, the discussion has focused on the diverse modes of colonialism and subsequent variation that existed across the institutional ideology of the colonisers at the time of conquest, and how their stance transformed with changes in technologies as colonial extraction intensified. However, to understand why such variation existed among the colonies under coloniser with similar institutional backgrounds, one needs to examine the key differences among the colonies taken under siege. In other words, it is not only the variation in the political economy of the metropolitan state that mattered but also the existing political and economic organisation of state being colonised at the time of the conquest and its institutional

¹⁷A quasi-feudal attitude toward the workforce in the phase of mercantilism kept the profitability on the high side for the early industrialists through various means (Mohaney, 2010). For example, profitability was kept high through biased policies within colonies in favour of commercial export agriculture (cash crops) against the subsistence food grain agriculture required for food security in colonies or through investments into the mining industry. Cheap slave labour was instrumental both because of its lower cost and also because the technologies at the time were not health-friendly (Fenoaltea, 1984; Fogel & Engerman, 1974). Hence, colonial extractive institutions under the technological frontier of the time when such practices took place, did play a vital part in the development of the global economy through creating favourable conditions for industrialisation.

¹⁸Colonial wealth directly helped in the process of industrialisation within Britain and Holland, not just financially but also in the form of cheap access to inputs such as labour, raw materials, and minerals (Stavrianos, 1981). In contrast, colonial extraction of Spain and Portugal indirectly supported the process of industrialisation not of themselves but other European countries like Britain, etc. since their population indulged in the consumption of industrial products produced elsewhere rather than their production processes at home.

¹⁹Similar school of thought in the sense that whether it is the phenomenon of settler or non-settler colonialism as emphasised in AJR description of history or distinction based on liberal or mercantilist paradigm as has come up in general literature.

complexity²⁰ also acted as a vital limiting factor for the path colonisers eventually took for their various colonies (Austin, 2008;²¹ Frankema, 2010; Mohaney, 2010).

Population density of the colonies at the time of conquest not only reflected their intensity of resistance as stressed by AJR but also proxied their degree of hierarchy within their pre-colonial institutions—which has shown to define stakes for the liberal and mercantilist colonisers differently in literature. Complex society with a dense population and socially divided setup were in accordance with the institutional structure of mercantilist rulers (Mohaney, 2010, Frankema, 2010, p. 424). While on the other hand, a thin population base with no pre-colonial social stratification would require inclusive institutional reforms by the mercantilist colonisers (Mohaney, 2010). Hence, in territories where such pre-colonial base was present, it had acted as a shield towards the extractive mode of colonisation for the mercantilist administrations.²²

Wherever colonisers with mercantilist orientation found already in place extractive pre-colonial structures i.e. slavery or huge tax base in the form of dense indigenous presence and institutional hierarchies, it gave them incentive to reinforce exploitative institutions. The congruence in such pre-colonial extractive institutions and mercantilist ideology had led to a deeper transgression in these societies.²³ On the other hand, for liberal colonisers, profit motives required participatory and market-oriented approach, but only where their immigrants have an edge of being settled in the new environment as a majority.

The stakes of liberal colonisers, hence, differed from coloniser with mercantilist ideology when facing thin versus dense surviving indigenous populations at the time of the conquest (Mohaney, 2010). Areas with sparse population had provided them with an opportunity to wipe out most of the indigenous population to overwrite the pre-colonial institutions with their free-market-oriented institutions through heavy settlements of their immigrants. On the other hand, dense surviving population with complex stratified pre-colonial institutions had acted as a disincentive for institutional spread in lines with liberal ideology since the cost of transforming the existing extractive institutions was

²⁰Institutional complexity of pre-colonial societies as proposed in Mohaney (2010) thesis as the crucial feature of the extent to which colonisers could transfer obvious elements of their institutional setup into the new territory needs to be assessed taking into consideration political, economic and cultural formation of such societies as a barometer. The move from least to most complex societies has been judged by an anthropologist by assessing how politically there is a move from decentralised and non-bureaucratic governance to a centralised authoritative structure with complex patrimonial-bureaucratic controlling mechanism, economically from relatively non-specialised production techniques without division of labour to organised agriculture with advanced specialisation of labour with the use of coercive methods and finally culturally from a small homogenous population to stratified large populations with exploitative practices in place in terms of ethnicity and other social divisions (Flannery, 1972). Accordingly, by above criteria, anthropologists have characterised the pre-colonial societies into three levels of institutional complexity: among the least complex are those who have termed as hunter-gathering, herding, and horticultural communities, under intermediate levels comes those that have been referred to in the literature as chiefdoms or advanced horticultural societies and finally among the most complex societies comes those named as “proto-states,” states or advanced civilisations (ibid).

²¹Austin (2008) explains constraints within which colonial powers had to work given how pre-colonial institutions arose in the face of factor endowment proportion facing a region—more specifically land abundance view in case of Africa.

²²This point is all the more relevant for mercantilist rule as it was structured around the exploitative mechanism and coercive labour as a tool for a profit-maximising rule.

²³To make the point, let us rephrase the point that we are emphasising here that mercantilist coloniser had a good fit with those societies that had surviving dense native communities with already existing social, political, and economic divisions than in colonies where natives and their institutions have been decimated and marginalised (Mohaney, 2010).

higher especially in the presence of hierarchical indigenous base. Moreover, the possibility of developing their settlement to dominate demographics in favour of their settlers against the natives also seemed unlikely with surviving dense communities of indigenous people after their colonial victory.

The ultimate direction in which colonies under colonial administration with the mercantilist or liberal outlook from their home institutional legacy moved, hence, depended on the constraints put by the size of surviving native population and their pre-colonial organisational structures. However, the size and structure of the native population had a varied impact on the degree to which mercantilists and liberal colonisers would install institutional features to their acquired colonial territories (Mohaney, 2010). Therefore, the process of colonialism and development can be best explained as an evolving fit between the institutions of the colonising nation and those of the colonised territory, and it is this interaction that holds the key to post-colonial outcomes, which cannot be accounted within AJR framework.

2.2b. Strategic Interplay of Various Stakeholders and Differential Institutional Course

According to this viewpoint, it is not just merely the initial conditions at the time of colonisation as stressed in AJR thesis²⁴ but also the relative bargaining strength of various stakeholders within the colonies that had led to varied institutional path across colonies. A neat example of such a role of native agency in the historical evolution under colonial rule comes from the varied experience of areas under the direct and indirect rule²⁵ of same colonial administration. Empirical evidence in Banerjee & Iyer (2005) and Iyer (2010) confirms that colonies under the indirect control of British administration where native autonomy was maintained outperformed regions that were under direct control in terms of contemporary economic outcomes.

A relatively higher level of governance and homogeneity among population in Indian regions that had a history of being under indirect British rule indicated how colonial history emerging from the relative bargaining power of the natives and colonial administration can set the trajectory for institutional evolution (Banerjee & Iyer, 2005; Iyer, 2010). However, dynamics may vary from case to case as there is evidence of non-uniform impact of indirect colonial rule on post-colonial outcome in different regions (Frankema, 2010). Hence, generalised conclusions as to how indirect versus direct rule may have provided natives' stakeholders' reasons to influence colonial policies with a forward-looking developmental agenda should be inferred with caution, as the experience may vary depending on the strength of natives and colonisers' resolve in implementing effective institutional organisations.

More enlightening perspective to differential institutional development across colonies regarding the strategic interplay of colonial and indigenous elements comes out

²⁴By initial conditions we mean settler mortality risks and native population density at the time of the conquest.

²⁵Why direct or indirect rule was chosen by Britain in different parts of India may have depended on a number of factors such as what kind of pre-colonial institutions were present at the time of colonisation (Gazdar, 2011), what were geographical constraints (Morrison, 2006 or Engerman & Sokoloff, 1997) and also the requirements of colonisers at time of conquest (Mohaney, 2010). Since this is not our key concern so we will not indulge in this discussion here.

in Frankema (2010).²⁶ In this study, author investigates why three colonies namely Northern Rhodesia, Sierra Leone and Malaysia under the indirect rule of same colonial power and with somewhat similar geographical conditions show different institutional trajectories. To elaborate further, the study explains different patterns of land distribution (egalitarian or stratified) in terms of the strategic interaction of the various stakeholders involved in such process: the colonial administration, colonial settlers and native population.

The important finding that comes out from Frankema (2010) is that even though all these colonies were administered under indirect colonial rule of same colonial power, yet different land ownership patterns evolved across these regions. Establishment of coercive institutions with relatively greater distribution of land from natives to white settlers appeared in Northern Rhodesia. In Sierra Leone, Europeans settled to a smaller extent despite evidence of its friendlier disease ecology and subsequent lesser mortality risk for the settlers.²⁷ Similarly, Malaysia, even though possessed very similar geographical and organisational structure in the pre-colonial rule as Northern Rhodesia and Sierra Leone but once under the indirect control of colonists showed the presence of higher land inequality than Sierra Leone. The prime reason for these divergent patterns as per analysis in Frankema (2010) was found in how natives were placed against colonial rule in each colony with highest degree of their relative strength in Sierra Leone and lowest in Northern Rhodesia. (Frankema, 2010).²⁸

Hence, in the above framework of Frankema (2010), the comparative evidence of three colonies with somewhat similar ecological conditions and hence similar disease environment for settler²⁹ becomes an effective tool to highlight the role of political economy aspects in the institutional evolution. Such an adjustment of indigenous voice and how colonists adapted their policies to keep social and political support for the smooth functioning of their rule cannot be feasible in AJR theoretical insight of exogenous imposition of extractive or inclusive institutional base in response to disease environment and native population density.

2.2c. Use of Indigenous Forces to Support Institutional Course

AJR's description of colonial history missed the role of pre-colonial institutions and how it impacted the institutional choice of the colonisers in due course of their colonial rule. So far, in the discussion above, we have highlighted how pre-existing

²⁶This evidence has importance because in its analysis, Frankema (2010) compares colonies that were under same colonial administration, all were under indirect rule and all had very similar geographical and pre-colonial organisational structures. Hence, author in this study is able to control all these factors while comparing different degrees of native agency relative to colonial hold within these colonies and its impact on land inequality. Most of other evidence is not able to control for these factors rather their evidence captures impact of differences in pre-colonial geographical factors on the divergent institutional ends such as in Morrison (2006) or Engerman & Sokoloff (1997). Hence, the evidence of Frankema (2010) is unique since it isolates the impact of strategic interplay between natives and colonisers from other plausible reasons- an analysis we did not find in other works on understanding colonialism.

²⁷This pattern of restricted European settlement despite minute mortality risk for the settlers in Sierra Leone provides evidence against the AJR premise.

²⁸The strength of colonial hold on Malaysia was relatively more than that in Sierra Leone but remained much limited compared to Northern Rhodesia, as can be inferred by the import of labour from India and China to work on Malaysian plantation than the use of native Malay population. This shows that colonists found it profitable to import labour to meet labour scarcity in Malaysian colony than bear risk of conflict with natives in case of imposing coercive practices of forcing the native population (Frankema, 2010).

²⁹To remind settler mortality is a vital source of identification in AJR research.

indigenous institutions interacted with the mercantilist and liberal orientation of the colonisers to produce limitations on the colonial institutional action. However, here we will follow the same discussion with a different angle.

Apart from evolving fit of pre-colonial social organisations and colonial institutional ideology, our focus in this section will be on how the forces emerging from within the indigenous pre-colonial setup had acted as defining factors for the colonists. To be specific, here the constraints imposed on colonist's institutional pursuit, especially in case of African history will be evaluated in the form of local will of native agency.³⁰ Further, British colonisation of Indian subcontinent will be used as a case study to highlight tangible limits in terms of local traditions and pre-colonial treaties and how their violation had the power to not only jeopardise the colonial hold but also affect their revenue collection objective adversely (Wilson, 2016).

As discussed in the previous section, colonial powers did use the extractive institution of slavery to their advantage to fill up the labour shortages at minimal prices in their countries or colonies elsewhere in terms of brutal intensive use of African slaves (Manning, 1990, Fenoaltea, 1999). Access to cheap labour for European powers in the form of slavery did act as a vital stimulant to their initial growth takeoff (Inikori, 2002). However, it is crucial to understand European hold in choosing the extraction rate of material and human trade in the African continent.

AJR conceptual formation becomes questionable once African colonial experience is evaluated taking colonialism in its true spirit of territorial control, which raise some interesting loopholes in AJR account of history. Indulgence in African slave trade by Europeans was documented as either an act of raiding as in initial European ventures or through formal dealing with African slave merchants (Dike, 1956). Slaves acquired through raids can be characterised as a crime or defiance of a country's sovereignty. While the slaves acquired through the act of purchase becomes a matter of trade between local African slave traders and their European counterparts (Iliffe, 1995; Fenoaltea, 1999).

Moreover, if European indulgence in the slave trade with Africa is considered as part and parcel of colonialism and its extractive institutional legacy according to AJR, then how can the existence of historical links between Arabs and African slave trade and its impact on institutional course for Africa can be accommodated within AJR reasoning? If such an act of buying African slaves by non-European countries has not been described as a form of colonialism in literature, then AJR notion of European extraction in Africa becomes a valid contradiction to historians' point of view on colonialism.

The historical evidence shows European colonial powers had no means to control slave trade rather their indulgence into African slavery was only through the ability of their traders to deal effectively with local African traders (Thornton, 1998: pp.100–101, pp. 114–116). Hence, the local pre-colonial institution of human pawning and the will of local African chiefdoms to sell their acquired slaves played an important role in the establishment of extractive institutional formation for African continent rather than the institutional transfer from the European colonial powers.³¹

³⁰Here it is important to note that in AJR framework, pre-colonial practices of slavery has been counted as part of the colonial extraction in Africa. Hence, under AJR application, it is assumed that European colonisers' had some sort of control on slave supply from Africa.

³¹We should rather use the term imperial powers here as colonialism in literal definition apply to conquered dominance, which occurred to much-limited degree and for a much shorter time in African continent than others.

Further, we do have evidence where local chief lords had restricted their slave trade in response to their local conditions as was seen in case of kingdoms of Benin and Kongo that were among the main initiators of trade in humans to Portuguese (Thornton, 1998, pp. 110–125). This suggests that Europeans were in no means to influence the rate of extraction whether politically, economically or militarily in most of the colonial history which run counter to AJR theme. Further, the AJR notion of higher settler mortality rate leading to higher extraction cannot readily be applied in this context as Europeans faced higher deaths was not merely through disease environment but also in consequences of higher African resistance (Thornton, 1998, pp. 36–40). Hence, in such cases, African resistance not only limited the European interference in terms of their extractive practices but also provided grounds for African agency to choose the rate of extraction for their region themselves wherever and whenever they willingly sold their people to Europeans.

Coming to the second part of the argument that besides the indigenous participation in choosing their institutional path as had been seen in the case of Africa, the local traditions and pre-colonial treaties also sometimes acted as a binding constraint. Let us explain this point in context of Ryotwari land settlement that was enacted simultaneously in different parts of British India such as Sindh, Madras, Bombay and Bengal presidencies; however the institutional impact in Sindh was totally in contrast to other three regions due to certain pre-colonial practices. The reason was institutional discrepancy across these regions resulted from how British administration defined Ryot³² differently in Sindh and other regions.

Given, in Sindh, there existed an intermediary namely Zamindars—the tribal landlords that were traditionally entitled to receive one-sixteenth of the crop as a tax called *laapo* for investing in cultivating land through irrigation and for protecting the *hari*—the actual cultivator of land in difficult times. Due to existence of already in place traditional revenue apportioning system, colonial administration in Sindh decided to deal directly with these zamindars and declared them as the Ryot under the implementation of the settlement rather than the Hari who despite being original cultivator was treated merely as a serf. This preserved the power structure in favour of large landlords in British Sindh. In contrast, in Madras, Bombay and Bengal, the property ownership rights were given to actual cultivators leading to the establishment of individual rights and homogeneous society (Gazdar, 2011, pp. 11-12).

Though, Ryotwari land settlement was enacted in all these regions, yet its institutional impact differed in terms of how such regions evolved the land distribution patterns in consequence of being awarded property rights claims differently due to pre-colonial initial conditions.³³ Therefore, not only has the indigenous will to accept

³²Ryot by definition means the cultivator of land.

³³Why different forms of pre-colonial institutions emerged in India? Was it because geography of these areas that guided different pre-colonial practices in Sindh than elsewhere or something more. Indeed, geography can play an important role in emergence of certain pre-colonial institutions. For example, Marx explains extractive pre-colonial taxation system in India as a natural course due to certain geographical factors (Kohn & Reddy, 2017). As per his analysis given India's agricultural productivity was linked to large-scale public works such as irrigation, hence oriental despotism was bound to emerge since only a centralised state could finance such an activity. However, given this research question is independent than what is being addressed here, hence we will not indulge into this debate here. Rather what we are trying to do here is to argue that this is the case where we indeed find evidence of the pre-colonial institutional setup leading the Britain to opt for empowering different groups. And hence institutional path taken by them in India should be at best considered as the evolving fit of pre-colonial and colonial institution.

colonial policies had acted as a limiting force for colonial institutional impact as was highlighted through African experience but also the pre-colonial traditional settlements among the indigenous populations defined the route colonists³⁴ took in their institutional choices in case of Indian subcontinent under British colonialism.

2.2d. Importance of Native Agency in Choosing their Historical Path

Post-colonial Exposure

AJR notion of the external imposition of institutions by European powers builds on observation of the inverse correlation between mortality rates of the European settlers and the establishment of European-style inclusive institutions. Accordingly, in densely populated colonies that provided a much better avenue for extractive practice³⁵, colonial administration chose rent-seeking instead of preferring the participatory institutions. Hence, above line of reasoning by AJR maintains emergence of institutions in colonies as a choice of the colonial governments given the conditions for European settlement, negating any role of the native historical agency under colonial rule in influencing their institutional path. Such a thesis is very strong for a historian perspective and can easily be nullified through historical anecdotes. Theoretically speaking, even in most extractive colonial rule, its mere continuation had required local support whether through support of corrupt locals or through the compromises with locals to avoid revolt.

Coming to the first point, among a few such channels that prove the historical agency of local people as identified in the previous sections are twofold. Firstly, in the context of how variation in pre-colonial institutional organisation defined the limits and incentives differently for settlers and non-settler colonialism within mercantilism and liberalism paradigms in the colonial exposure (Mohaney, 2010). And secondly, in terms of how pre-colonial treaties or social organisation had made colonists adjust their system of property evaluation for revenue collection differently (Gazdar, 2011). But more importantly, for our argument here is to go beyond above channels to look at how colonisers looked for a match with pre-existing exploitative social structures to identify corrupt native elites as a way to strengthen their hold (Mohaney, 2010). Therefore, local political will to submit to colonialism also came from forces within the colonised region at some level which needs due recognition.

Also just as a will to submit to colonial rule was present in some segments of population, so was resistance against such rule which has an important role in protecting their natives business and properties too. For example, the influence of native's resilience can be seen through survival of native production in Southern Rhodesia, and Kenya despite the colonial pressure of pushing the labour out of their own production into the labour market or the emergence of world's most prominent cocoa belt in Ghana during its colonial subjugation served as an example of how indigenous forces organised for interests against colonial extraction (Mosley, 1983; Hill, 1963). Similarly, the establishment of Northern Nigeria under British colonial rule as a prime exporter of peanuts and not of raw cotton as was intended under colonial enterprise to support their textile industry, again present the case of native agency (Hogendorn, 1979).

³⁴Account of British colonialism of India in Wilson (2016) show that the colonial administration did start at to replace local negotiation and mutual obligations between landlords, peasants, and the officials at some point but that led to conflicts and riots destabilising their control on India. Wilson (2016, p. 80., p. 101, p. 233 and p. 293).

³⁵By extractive practice what is implied here is forced labour or higher tax rates.

Decimation of local's industries to avoid competition was a common practice within British colonialism, for example, consider the case of marginalisation of textile industry in Egypt and India (Stavrianos, 1981, Bagchi, 1984, p. 82 and Tharoor, 2017). Further, in case of India, analysis of how local Indians took back almost all businesses and industries from the expatriate companies in Eastern India by 1950 show that this process did not happen overnight. Rather it was case of proactive native resistance that had started around the time of First World War (Ray, 1979, Tomlinson, 1981). Hence, it will be not wrong to infer that wherever the natives had kept their production processes in their control under colonial rule does indeed reflect their coordinated agency in protecting their businesses from expropriation of foreign companies as is evident from colonial experiences of Southern Rhodesia, Kenya, Ghana and Northern Nigeria (Mosley, 1983; Hill, 1963; Hogendorn, 1979, Austin, 2005, 2007).

Therefore, the point that needs its right credit from a historian perspective and which has been hidden within AJR methodological simplification of a complex historical issue of colonialism and its impact is that even in most exploitative colonial rule, the indigenous support forces at one level had acted as a backbone to the feasibility of that imposed external administration and at other colonial policies could be slow but eventually did adjust to seek general political acceptance of the indigenous masses for their rule (Tomlinson, 1981).³⁶ For example, the property rights mechanism that was established within the colonial rule for whatever colonial objectives of tax base records, did respond to limits set by local opposition to changes. A case to this points to the differential imposition of land rights in Sindh and Bengal within British India following the local will of powerful groups (Banerjee and Iyer, 2005; Gazdar, 2011, pp. 11-12).

Hence, it is quite feasible that colonists who intended to establish certain institutions which reflected their setup in the metropolitan base could not implement or if implemented were reversed in congruence with past pre-colonial legacies and existing mistrust of indigenous masses for the colonial administration culminating into organised resistance from natives against the colonial hold (Wilson, 2016). Example of this can be found in Mutiny of 1857 in British India which was evidence of local resistance to British administration that became a coordinated force by 1919 or continuation of indigenous property rights systems as done in some African colonies cited above in face of local support for such a system. Hence, colonial administration could be slow but eventually found in their interest to compromise keeping in view political and economic will of their subjects, and such examples of a shift in colonial policies to adjust the indigenous agency blur the AJR implication of external institutional transfer dependent only on colonial power's motives based on settlement criterion.

³⁶The expatriate businesses of Eastern India that were operating as monopsony in context of internal trade and dominating most of foreign trade from the region till 1900 lost most of their businesses to local Indians by 1950's as Indian support for the colonial rule started to diminish (Bagchi, 1971). The analysis in Tomlinson (1981), show that many factors played their due role. But one thing that was most dominant included how the political uncertainty that followed post increased native resistance against the colonial rule de-incentivised these expatriate companies to limit their expansion into new fields such as sugar and cement. This not only opened avenue for native entrepreneur to fill the gaps but also build up finances to later expand in other industries such as paper, chemicals and machine tools etc.

3. CONCLUSION

Our critique on Acemoglu, Johnson, and Robinson's (2001, 2002) does not undermine the significance of their findings in any way. We, in our research, maintain that the ingenuity of AJR research lies in its use of unique instrumental variable found through the impact of settler mortality rates on the kind of institutions they placed which could not have been possibly affected by current growth outcomes. Hence, in the above context, not only their grasp on econometric tools is admirable but also their explanation as to how growth-enhancing institutions appeared in some parts of the world and failed to emerge in others is of crucial importance.

However, in our analysis, we show that AJR assumption of exogenous imposition of colonial institutions based on conditions for their settlement is too strong for two reasons. First, one account for the impact of Atlantic slave trade on colonial metropolitan states in terms of technological frontier and how home institutions evolved, it brings forth first source of endogeneity. That is how such dynamics had defined the incentives for mercantilist and progressive colonialism differently for European colonial powers especially the ones that have embarked on industrialisation. And secondly, we find that imposed colonial institutions on the colonies can only be best described as an evolving fit between colonial and pre-colonial institutions with a defining role of native agency in sustainability of the final outcome. Hence, some degree of endogeneity did exist where conditions in colonies other than those related to settlement for the colonisers had indeed played an important role in defining the institutional path taken by colonial powers within their colonies.

Our findings have meaningful implication in the context of how foreign aid often termed as means to impose neo colonialism/imperialism functions in developing countries. Just as in case of our analysis, pre-colonial institutions set-up of colonies defined the trajectory of colonial rule, in case of foreign aid the institutional weaknesses of the developing countries which often are rooted in their colonial past serve as the cause of aid inflows in its first stage and as source of further weakening of their institutions in its second phase. This creates a recurrent need to look out for aid and loans in such developing countries undermining their long-term future growth prospects.

REFERENCES

- Acemoglu, D., Johnson, S., & Robinson, J. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5), 1369–1401.
- Acemoglu, D., Johnson, S., & Robinson, J. (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics*, 117(4), 1231–1294.
- Acemoglu, D., Johnson, S., & Robinson, J. (2005). The rise of Europe: Atlantic trade, institutional change, and economic growth. *The American Economic Review*, 95 (3), 546–579.
- Acemoglu, D., Johnson, S., & Robinson, J. (2012). The colonial origins of comparative development: An empirical investigation: Reply. *American Economic Review*, 102(6), 3077–3110.

- Albouy, D. (2004a). *The colonial origins of comparative development: A reinvestigation of the data*. Berkeley: University of California.
- Albouy, D. (2004b). *The colonial origins of comparative development: A reexamination based on improved settler mortality data*. Berkeley: University of California.
- Albouy, D. Y. (2012). The colonial origins of comparative development: An empirical investigation: Comment. *American Economic Review*, 102(6), 3059–76.
- Altaf, S. W. (2011). *So much aid, so little development: Stories from Pakistan*. Woodrow Wilson Center Press.
- Amin, S. (2018). *Modern imperialism, monopoly finance capital, and Marx's law of value*. New York: Monthly Review Press. 131–35.
- Angeles, L. (2007). Income inequality and colonialism. *European Economic Review*, Elsevier, 51(5), 1155–1176.
- Austin, G. (2005). *Labour, land, and capital in Ghana: From slavery to free labour in Asante, 1807-1956*. Rochester, NY: University of Rochester Press.
- Austin, G. (2007). Labour and land in Ghana, 1874-1939: A shifting ratio and an institutional revolution. *Australian Economic History Review*, 47(1), 95–120.
- Austin, G. (2008). Resources, techniques, and strategies south of the Sahara: Revising the factor endowments perspective on African economic development, 1500-2000. *Economic History Review*, 61(3), 587–624.
- Bagchi, A. K. (1971). *Private investment in India 1900-1939*. p. 166. Cambridge University Press.
- Bagchi, A. (1984). *The political economy of underdevelopment*. Cambridge: Cambridge University Press.
- Bairoch, P. (1993). *Economics and world history: Myths and paradoxes*. University of Chicago Press.
- Banerjee, A. V., & Iyer, L. (2005). History, institutions, and economic performance: The legacy of colonial land tenure systems in India. *The American Economic Review*, 95(4), 1190–1213.
- Bardhan, P. (2005). Institutions matter, but which ones? *The Economics of Transition*, 13(3), 499–532.
- Bertocchia, G., & Canova, F. (2002). Did colonisation matter for growth? An empirical exploration into the historical causes of Africa's underdevelopment. *European Economic Review*, 46, 1851–1871.
- Bishai, L. S. (2004). Liberal empire. *Journal of International Relations and Development*, 7(1), 48–72.
- Chakraborty, A. (2017). Colonial origins and comparative development: Institutions matter.
- Coleman, D. C. (1957). Eli Heckscher and the idea of mercantilism. *Scandinavian Economic History Review*, 5(1), 3–25.
- Crow, B., & Thorpe, M. (1988). *Survival and change in the third world*. New York: Oxford University Press.
- Diamond, J. M. (1999). *Guns, germs & steel: The fate of human societies*. New York: W. W. Norton and Company.
- Digby, W. (1969). *"Prosperous" India: A revelation from official records*. New Delhi: Sagar Publications.

- Dike, K. O. (1956). *Trade and politics in the Niger delta, 1830–1885*. Oxford: Oxford University Press.
- Engerman, S. L., & Sokoloff, K. L. (1997). Factor endowments, institutions, and differential paths of growth among new world economies: A view from economic historians of the United States. In Harber, S. (ed.) *How Latin America fell behind* (pp. 260–304). Stanford: Stanford University Press.
- Engerman, S. L., & Sokoloff, K. L. (2005a). Institutional and non-institutional explanations of economic differences. In Menard, C. and Shirley, M. M. (eds.), *Handbook of New Institutional Economics* (pp. 639–665). Dordrecht, Springer.
- Engerman, S. L., & Sokoloff, K. L. (2005b). The evolution of suffrage institutions in the Americas. *Journal of Economic History*, 65, 891–921.
- Ekelund Jr., R. B., & Tollison, R. D. (1980a). Mercantilist origins of the corporation. *Bell Journal of Economics*, 11(2), 715–720. The Rand Corporation.
- Ekelund Jr., R. B., & Tollison, R. D. (1980b). Economic regulation in mercantile England: Heckscher revisited. *Economic Inquiry*, 18(4), 567–599. Western Economic Association International.
- Easterly, W. & Pfitze, T. (2008). Where does the money go? Best and worst practices in foreign aid. *Journal of Economic Perspectives*, 22(2), 29–52.
- Easterly, W. & Levine, R. (2016). The European origins of economic development. *Journal of Economic Growth*, 21(3), 225–257.
- Fenoaltea, S. (1984). Slavery and supervision in comparative perspective: A model. *The Journal of Economic History*, 44(3), 635–668.
- Fenoaltea, S. (1999). Europe in the African mirror: The slave trade and the rise of feudalism. *Rivista di storia economica*, 2, 123–166. Societa Editrice il Mulino.
- Feyrer, J. & Sacerdote, B. (2009). Colonialism and modern income: Islands as natural experiments. *The Review of Economics and Statistics*, 91(2), 245–262.
- Flannery, K.V. (1972). The cultural evolution of civilisations. *Annual Review of Ecology and Systematics*, 3, 399–426.
- Fogel, R. W., & Engerman, S. L. (1974). *Time on the cross: Evidence and methods—A supplement*. Boston: Little, Brown and Company.
- Frankema, E. (2010). The colonial roots of land inequality: Geography, factor endowments, or institutions? *The Economic History Review*, 63 (2), 418–451.
- Gazdar, H. (2011). The fourth round, and why they fight on: An essay on the history of land and reform in Pakistan In *Levelling the playing field: A survey of Pakistan's land reform*, PANOS South Asia, Kathmandu, pp. 8–65.
- Grier, R. (1999). Colonial legacies and economic growth. *Public Choice*, 98(3/4), 317–335.
- Hall, R. E., & Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics*, 114(1), 83–116.
- Haq, N. U. (1999). Reform efforts in Pakistan: Limited success and possibilities for the future. In Khan, S. R. (ed.) *Fifty years of Pakistan's economy traditional topics and contemporary concerns*. Oxford: Oxford University Press, pp. 29–50.
- Haq, N. U. (2018a). *Looking back: How Pakistan became an Asian tiger in 2050*. Kitab (Pvt) Ltd.

- Haque, N. U., Mahmood, U., Abbas, S., Lodhi, A., Rab, M., & Jones, C. S. (2018). *The university research system in Pakistan* (British Council Knowledge Platform Project Report).
- Harvey, D. (2003) *The new imperialism*. Oxford: Oxford University Press.
- Hayter, T. (1971) *Aid as imperialism*. Harmondsworth: Penguin.
- Hill, P. (1963). *The migrant cocoa-farmers of Southern Ghana: A study in rural capitalism*. Cambridge: Cambridge University Press.
- Hogendorn, J. S. (1979). *Nigerian groundnut exports: Origins and early development*. Zaira, Nigeria: Ahmado Bello University Press.
- Hundt, D., & Bleiker, R. (2007). Reconciling colonial memories in Korea and Japan. *Asian Perspective*, 31(1), 61–91.
- Ikenberry, G. J. (2006). *Liberal order and imperial ambition: Essays on American power and international order*. Polity.
- Iliffe, J. (1995). *Africans: The history of a continent*. Cambridge, UK: Cambridge University Press.
- Inikori, J. E. (2002). *Africans and the industrial revolution in England*. Cambridge University Press.
- Iyer, L. (2010). Direct versus indirect colonial rule in India: Long-term consequences. *Review of Economics and Statistics*, 92(4), 693–713.
- Keefer, P., & Knack, S. (1997). Why don't poor countries catch up? A cross-national test of an institutional explanation. *Economic Inquiry*, 35(3), 590–602.
- Kiely, R. (2005). *Empire in the age of globalisation: US hegemony and neoliberal disorder*. Pluto Press.
- Kohn, M. & Reddy, K. (2017) "Colonialism", The Stanford Encyclopedia of Philosophy (Fall 2017 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2017/entries/colonialism/>>.
- Kohli, A. (2004). Front matter. In *state-directed development: Political power and industrialisation in the global periphery* (pp. I-Vi). Cambridge: Cambridge University Press.
- Lenin, V. (1917). *Imperialism, the highest stage of capitalism*. (Edition 1948). London: Lawrence and Wishart.
- Lumumba-Kasongo, T. (2011). China-Africa relations: A neo-imperialism or a neo-colonialism? A reflection. *African and Asian Studies*, 10(2-3), 234–266.
- Mabee, B. (2004). Discourses of empire: The US 'empire', globalisation and international relations. *Third World Quarterly*, 25(8), 1359–1378.
- Mahoney, J. (2010). *Colonialism and postcolonial development: Spanish America in comparative perspective*. Cambridge University Press.
- Manning, P. (1990). *Slavery and African life*. Cambridge, UK: Cambridge University Press.
- Marx, K. (1976) *Capital vol I*. (London: Penguin, 1976), 27–28.
- Melleuish, G. (2001). A short history of Australian liberalism. (CIS Occasional Papers, 74).
- Mosley, P. (1983). *The settler economies: Studies in the economic history of Kenya and Southern Rhodesia, 1900–63*. Cambridge: Cambridge University Press.

- Morrison, M. (2006). Ghana's political parties. How ethno-regional variations sustain the national two-party system. *Journal of Modern African Studies*, 44(4), 623–647.
- Nunn, N. (2008a). The long-term effects of Africa's slave trades. *Quarterly Journal of Economics*, 123, 139–176.
- Nunn, N. (2008b). Slavery, inequality, and economic development in the Americas: An examination of the Engerman-Sokoloff hypothesis. In Helpman, E. (ed.) *Institutions and economic performance*. Cambridge: Harvard University Press, 48–180.
- O'Rourke, K. & Williamson, J. G. (1999). *Globalisation and history: The evolution of a nineteenth-century Atlantic economy*. Cambridge, Mass.
- Quinn, A. & Michael C. (2007). For better, for worse: How America's foreign policy became wedded to liberal universalism. *Global Society*, 21(4), 499–519.
- Ray, R. K. (1979). Industrialisation in India: Growth and conflict in the corporate private sector 1914–47. pp. 356–8.
- Reno, W. (1995). *Corruption and state politics in Sierra Leone*. New York: Cambridge University Press.
- Rodrik, D., & Subramanian, A. (2003). The primacy of institutions. *Finance and Development*, 40(2), 31–34.
- Rogowski, R. and Duncan, M. (2004). Inequality and Institutions? What Theory, History, and (Some) Data Tell Us. Paper presented at American Political Science Association Meeting, Chicago, September, 2004.
- Sang, L. W. (2009) *Collaborative colonial power: The making of the Hong Kong Chinese*. Hong Kong: HKU Press.
- Sang, L. W. (2018). Colonial modernity and the predicament of liberalism in Hong Kong. *Hong Kong and Liberty Series*, 11(26).
- Sang, L. W. (2018). Colonial Modernity and the predicament of liberalism in Hong Kong. *Global-E*. 11(26).
- Stavrianos, L. S. (1981). *Global rift: The third world comes of age*. New York: William Morrow & Co.
- Stephanie I. (2018). Foreign aid a facilitator of neo-Colonialism in Africa?
- Sylwester, K. (2008). Historical legacy and institutions across countries. *Journal of Applied Economics*, 11(2), 373–398.
- Thornton, J. (1998). *Africa and Africans in the making of the Atlantic World, 1400-1800*. Cambridge University Press.
- Tomlinson, B. R. (1981). Colonial firms and the decline of colonialism in Eastern India 1914–47. *Modern Asian Studies*, 15 (3), 455–486.
- Warren, B. (1973). Imperialism and capitalist industrialisation. *New Left Review*, 1/81.
- Williams, D. (2000). Aid and sovereignty: Quasi-states and the international financial institutions. *Review of International Studies* 26(4), 557–573.
- Williams, D. (2008). *The World Bank and social transformation in international politics: Liberalism, governance, and sovereignty*. London: Routledge.
- Williams, D. (2018). Liberalism, colonialism and liberal imperialism. *East Central Europe*, 45(1), 94–118.

Friendly Fire: Wheat Subsidy in Punjab, Pakistan

MUHAMMAD AHSAN RANA and MUHAMMAD NADEEM MALIK

During the past decade every year the Punjab government has faced an awkward situation at wheat harvest. It must buy millions of tons of wheat at an above-market price despite massive carry-forward stocks already lying in its granaries. Cost of procuring such huge quantities and subsidising sales to flour mills is enormous and is met with commercial borrowing. In recent years, the government has struggled to balance its accounts for wheat operations; its outstanding liability to commercial banks stood at Rs. 444.7 billion in June 2018, viz. 22 percent of the total budget of the province in 2017-18. Clearly, the government procures more wheat annually than it needs with borrowed money that it cannot pay back. Ostensibly, wheat procurement aims to benefit small farmers, but its procedures exclude, rather than include them. Conversely, consumers end up buying expensive wheat-flour. A significant beneficiary of the procurement regime appears to be commercial banks finance the procurement and earn interest thereupon.

Keywords: Food Subsidy, Wheat Procurement, Subsidy Reform

1. INTRODUCTION

Several recent studies (Aamir, 2017; Ahmad and Farooq 2010; Amid 2006; Cummings, et al. 2006; Dorosh and Salam 2008; Prikhodko and Zrilyi 2013; Scott and Hernandez 2018; World Bank, 2010) have highlighted inefficiency, high cost and wastefulness of food subsidy programs in developing countries. These programs usually started as small interventions with a specific mandate responding to a particular crisis (e.g. ration cards during WWII in several British colonies (Alderman, 1988), but they gradually morphed into large generalised subsidy regimes that by the end of the 20th century were significantly contributing to what O'Connor (1979) and Offe (1984) call 'the fiscal crisis of the welfare state.' Yet, subsidy rollback—even of the obviously wasteful types—has been difficult, and governments have struggled to design and implement reform packages that reduce costs without unduly harsh political backlash (Gutner, 2002).

Often these subsidies are justified with reference to the need to support farmer incomes and poor urban consumers (e.g. PFD Overview, n.d.). Since these are large groups, policymakers consider as politically unfeasible any proposal for a drastic

Muhammad Ahsan Rana <ahsan.rana@lums.edu.pk> is Associate Professor at the Suleman Dawood School of Business, Lahore University of Management Sciences, Lahore. Muhammad Nadeem Malik <malikn@unimelb.edu.au> is Senior Lecturer, Development Studies, School of Social and Political Sciences, The University of Melbourne, Australia.

reduction in food subsidy programs, their increasing costs and apparent ineffectiveness notwithstanding (Gutner, 2002). The untested assumption in these assertions is that food subsidies actually benefit farmers and poor urban consumers. This is an empirical question and must be settled as such.

Punjab's wheat subsidy programme—worth Rs. 30 billion a year (US\$ 216 million)¹—presents itself as a good candidate for deconstructing the claim that food subsidies benefit farmers and consumers. The wheat subsidy is Punjab's largest subsidy and a cornerstone of its efforts to promote food security (*Dawn* 15th April, 2017). Each year, the Punjab government procures millions of tons of wheat during harvest season at above-market prices and sells wheat to flour mills over the year at below-market prices. The objective is to protect farmers from a price crash at harvest time and to subsidise consumers' flour purchases (PFD Market Stabilisation, n.d.). During the past ten years, existing stocks and fresh procurement have comprised a massive inventory, much larger than government's average annual release to flour mills. To finance procurement operations, the government borrows money from various banks, but it has consistently failed to pay back its outstanding debt to banks and has been paying huge sums as the mark-up. Clearly, the Punjab government is procuring more wheat than it needs with borrowed money that it cannot pay back.

This paper is a critical examination of the wheat subsidy regime in Punjab. The Punjab Food Department (PFD) borrowed Rs. 115.4 billion from a consortium of banks to finance a massive procurement of 3.62 million tons in 2018. PFD procured this huge quantity despite a carry-forward stock of 3.59 million tons already lying in its granaries from previous year's procurement. Given that PFD's average annual release to flour mills (3-3.5 million tons) was smaller than its carry-forward stocks, there was no reason to procure more wheat in 2018. *Prima facie*, such massive procurement is done each year 'for the economic well-being of small farmers' (*Dawn* 15th April, 2017). But does it really benefit them, or are there other groups benefiting from this massive public intervention carried out in the name of small farmers?

The analysis below shows that small farmers are largely excluded from the process and receive only a fraction of the benefit, if at all. Furthermore, PFD is unable to pay back the money it borrows every year. Its outstanding liability to commercial banks stood at Rs. 444.7 billion in June 2018. This appears a case of 'friendly fire'—borrowing from military terminology whereby a military is hit by its own shelling (Sen, 2005: 212-15). Punjab's wheat subsidy is an example of a public intervention that ends up hurting the interests of those very poor whose interests it was supposed to protect.

This paper fills an important gap. Although, there is a large body of literature on food subsidies in developing countries, there are only a few published studies on Pakistan. No study dissects Pakistan's wheat subsidy to see where exactly the money is spent. Most studies (e.g. Alonso and Swinnen, 2016; Dorosh and Salam, 2008) focus on estimating how the benefit is shared between producers, consumers and intermediaries. Further, there is rarely an effort to distinguish between the interests of small, medium and large farmers, or to identify the institutionalised mechanisms that include/exclude various producer groups (see for example Aamir, 2017; Ahmad and Farooq, 2010). This paper, on the other hand, closely looks at official data to identify commercial banks as

¹US\$ 1 = Rs. 139 (in January 2019).

significant beneficiaries of wheat subsidy in Punjab. This is a surprising finding. The interest collection by banks on government's current borrowing and outstanding dues comprises $3/4^{\text{th}}$ (in some years > 90 percent) of the total amount brandished as wheat subsidy. This had escaped researchers' eye so far.

In addition to this short-term buildup of stocks and debt, this paper identifies the mechanisms that exclude small farmers from participating in official procurement campaigns as a more serious structural problem. By closely examining the grossly under-explored micro-processes of wheat procurement and the logistical life behind procurement numbers in Punjab, this paper invites and provokes policy communities to rethink the subsidy regime in Punjab so that a greater proportion of benefits reaches small farmers and urban consumers, in whose name the subsidy is provided.

2. METHODOLOGY

The case study method was considered appropriate to gain insights into considerations behind official decisions, their implementation and impact on identifying micro-processes that include/exclude various groups (Babbie, 2012). Secondary and primary data on wheat operations were collected during Feb-July 2017 and April-May 2018. Secondary data sources included PFD records and periodic surveys of the Pakistan/Punjab Bureau of Statistics. PFD records are public data since they are part of the (unclassified) record maintained by government departments in the normal course of their working. Anyone can obtain a copy upon a formal request under the Punjab Transparency and Right to Information Act, 2013. The authors were able to view these records after explaining to the Secretary PFD that the data were required for academic analysis. Authors' previous engagement with these officials for other policy research helped establish the *bona fide* of data collection. Data from the Pakistan/Punjab Bureau of Statistics were available on their websites. The analysis below uses PFD data for the past ten years only. Although, PFD maintains time series data for earlier period as well, but the same are not relevant to this analysis. Prior to 2008, there was no buildup of stocks² or outstanding bank debts.

Primary data were collected in the districts of Sargodha, Sheikhpura and Vehari—these are located in Western, Central and Southern Punjab respectively. Unstructured interviews were conducted with 35 respondents, including 15 farmers (eight had farms < 12.5 acres, five had farms between $12.5 - 25$ acres, and two had farms > 25 acres),³ nine PFD officials, four grain merchants, two flour miller, two urban consumers and three rural non-farm households. Initially, 45 farmers were selected from PFD's list of farmers (discussed later) in the field districts. But we managed to interview only 15 of them as the remaining were either absentee landlords, were out of town during our field work, or were simply unwilling to be interviewed. In PFD, we were able to interview two senior managers, one District Food Officer, and six officials working at Procurement

²PFD's time series data show that its carry-forward stocks have historically hovered around 0.25 million ton. They crossed one million ton in 2001, stayed the same in 2002, but declined quickly thereafter to levels below the usual. They rose again in 2006 to 1.4 million ton and declined thereafter. Consistent build-up of stocks to alarming levels is a recent phenomenon.

³These are called small, medium and large farmers respectively in this paper. Farmers of less than one acre are called very small farmers.

Centres. Other respondents were selected for convenience, i.e. people who were willing to talk to us. Each interview was conducted in person by the authors in Urdu for 50 minutes on average (the shortest took 25 minutes and the longest was slightly less than two hours). There was no opportunity to record interviews, so we took notes.

Our fieldwork coincided with procurement campaigns in 2017 and 2018 (i.e. mid-April to end May in both years), hence several interviews could be conducted *in situ*. We frequently visited key sites, such as procurement Centres, grain markets, bank branches, storage places, and spent time in villages observing the procurement as it happened. On different days, we spent a total of 28 hours in procurement centres, five hours in grain markets, four hours in banks, and 20 hours at various farms to carry out field observations. In addition to capturing stakeholder perspectives, this provided an opportunity to observe what Krishnamurthy (2012: 74) called the ‘micro-practices of procurement.’ We took field notes, which were later expanded into detailed observations to inform analysis and writing.

Further, one author had personal exposure to wheat operations through his association with an NGO (Fountain House) that has a wheat-producing farm (around nine acres) in Sheikhpura district for agro-based therapy of people with mental health conditions. The farm was able to sell its wheat to PFD in 2016, but not in 2017 and 2018 despite several follow up visits to concerned offices. This provided a first-hand experience of exclusion by the bureaucratic apparatus that ostensibly existed to benefit small farms like this.

The spread of fieldwork over two years and ethnographic observation of various actors provided critical insights into PFD’s wheat procurement process that would have been difficult to gain otherwise. Most of the fieldwork was carried out in 2017. However, official data on procurement and bank borrowing were updated in 2018, and short visits to (mostly the same) respondents and key sites enabled verification of observations across years and also allowed us to clarify confusions that had emerged while writing this paper.

3. PROCUREMENT TRENDS AND PRICE

Government’s intervention in wheat market dates back to 1968, when it procured wheat for the first time in response to bumper crop for two consecutive years that the market did not clear (Aziz, 1979). Gradually, the government became the principal buyer and the private sector’s role shrank. To keep flour prices low for poor households, it established an extensive network of ration-shops, which provided subsidised wheat flour to low-income households. Ration system was abolished in 1987 due to partial targeting, inefficiencies and corruption (Alderman, 1988; Islam and Garrett, 1997). The ration system was replaced with a subsidy on wheat issued to flour mills by the government from its procured stocks. Thus, a targeted subsidy was replaced by a general subsidy that ultimately became far more expensive than the one it replaced (Alderman, 1988). The twin requirements of clearing stocks during harvest season and providing subsidised wheat to flour mills later in the year led the government to procure progressively larger volumes of wheat each year.

As shown in Table 1, in the recent past, PFD has procured very large and variable quantities of wheat each year without an apparent link to production, existing stocks, or flour mills’ demand. On average, wheat procurement has been 18.6 percent of total

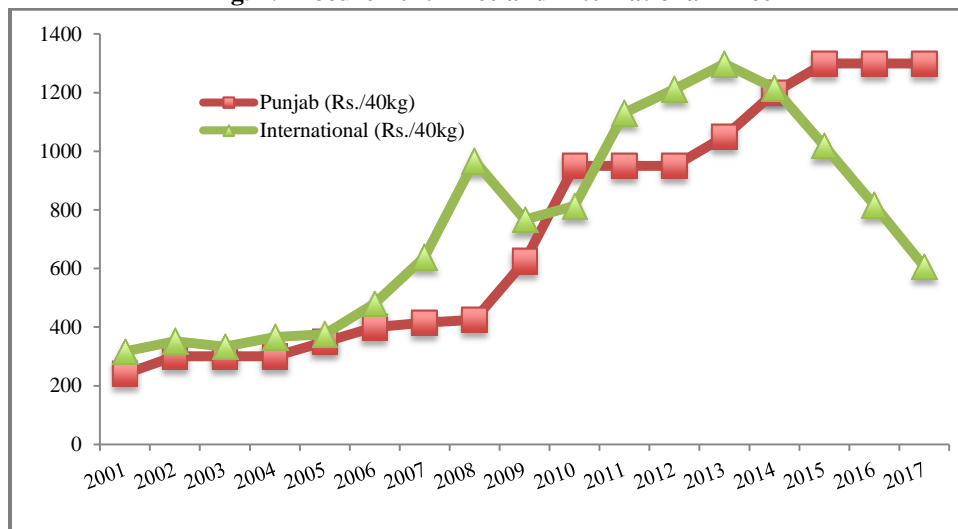
Table 1
Production and Procurement (Million Ton)

	Production	Existing Stocks	Procurement (% of Production)	Import (Export)	Total Available	Issued to flour Mills (% of Total Available)
2008	18.23	0.01	2.56 (14.0)	0.61 (0)	3.18	2.72 (85.5)
2009	18.42	0.23	5.78 (31.4)	0 (0)	6.01	3.04 (50.6)
2010	17.92	2.93	3.72 (20.8)	0 (0.91)	5.74	3.08 (53.7)
2011	19.04	1.88	3.19 (16.8)	0 (0)	5.07	3.20 (63.1)
2012	17.74	1.74	2.78 (15.7)	0 (0)	4.52	4.02 (88.9)
2013	18.63	0.48	3.68 (19.8)	0 (0)	4.16	4.02 (96.6)
2014	19.70	0.13	3.74 (19.0)	0 (0)	3.87	1.83 (47.3)
2015	19.20	1.96	3.23 (16.8)	0 (0.25)	4.94	2.90 (58.7)
2016	19.40	2.26	3.93 (20.3)	0 (0.44)	5.75	3.52 (61.2)
2017	20.46	2.58	3.95 (19.3)	0 (0)	6.53	3.60 (55.1)
2018	19.60	3.59	3.62 (18.5)	—	7.21	—

Source: Constructed from PFD data.

production. PFD has issued only a part of its available stock to flour mills and the rest has stayed in its granaries each year. Thus, in most years, PFD has significant carry-forward stocks from previous years. The peak procurement (5.78 million tons) during the past ten years was in 2009 and resulted in 2.93 million tons carry forward stocks at the time when procurement began in 2010. This must have translated into huge storage costs for PFD in 2009-10. Only small quantities of wheat are imported and exported, if at all. Mostly, import comprises food aid received from abroad. Small volumes of export despite massive surpluses can be explained in terms of the gap between the domestic and international prices, the former being higher than the latter (Aamir, 2017).

The procurement price-set in October-November each year —is supposedly based on the cost of production but is in practice determined by an interplay of technical and political considerations. Its comparison with the international price of wheat since 2001 (Figure 1) demonstrates that in most years the former has been below the latter. Procurement price has never decreased from its previous level: it has either increased from year to year or stayed the same. There were large price increases in 2008-09 and 2009-10 that appear to be delayed responses to international price escalations. No such response followed a decline in the international price since 2014. Clearly, there is no effort to peg the domestic price with the international price and to use international trade to stabilise prices and/or meet domestic demand.

Fig. 1. Procurement Price and International Price

Source: Constructed from International Grain Council and PFD Data.

Note: International prices were converted to PKR using currency conversion rates for 30th June in each year.

4. MICRO-PROCESSES OF PROCUREMENT

Wheat procurement is a tedious process in multiple ways, especially, for the small farmer in whose name such large procurement is justified. Several authorisations by petty officials characterise the process, which the farmer has to manage at his cost and time. This administrative burden is too heavy for most small farmers to carry and the process becomes the principal mechanism of their exclusion.

Tenants and small farmers often fail to cross the very first hurdle—their names do not appear on *patwari's*⁴ list, which is the basis of PFD procurement. Every year, when the procurement season commences, PFD sets up several procurement centres to which various villages are assigned. Each Centre has a list of farmers in its catchment area. This list has several errors. Tenants are often not included in the *patwari's* records on the insistence of influential landowners.⁵ Technically, a tenant can approach the PFD with a copy of the tenancy agreement to get his name included in PFD's list. But often tenancies are verbal, and few small tenants will prepare formal agreements⁶ just to satisfy PFD requirements.

Similarly, it is not uncommon for the names of politically weak or misaligned farmers to *accidentally* slip out of the *patwari's* list. Although, an aggrieved landowner whose name is not on *patwari's* list can approach higher officials for rectification. Our field work shows that this process is seldom initiated due to substantial transaction costs, such as multiple visits to various offices. Three out of eight small farmers in our sample and the Fountain House did not appear in the list in 2017. Interestingly, Fountain House

⁴ An important field official of the Punjab government.

⁵ A tenant recorded in *patwari's* records has certain protections against eviction.

⁶ To have a written agreement, a tenant needs willingness of the landlord to sign the agreement, to do the paperwork and pay the legal fees (e.g. stamp duty). After this is done and the tenant has his name on PFD's list, there is still no certainty that PFD would be procuring when his turn comes.

was on the list in 2016 when carry-forward stocks were small and PFD was procuring generously, but not in 2017 and 2018 when existing stocks were large. In-charge of Fountain House told us that his staff had visited their procurement centre each year to find out why their name was not on the list. Each time, they were returned with an advice to apply to the local Assistant Commissioner. The In-charge weighed the pros and cons of doing so, and in the end decided to sell to the local *beopari* instead.

Farmers whose names are in *patwari*'s list proceed to the procurement centre to obtain jute/polypropylene bags (called *bardana*). Farmers can supply their wheat only in officially-issued *bardana*. Each bag bears a number which is recorded against the farmer's name. Farmer's *bardana* allowance is based on the national average yield of 800 kg per acre (Pakistan, 2016), i.e. eight jute bags (100 kg each) or 16 polypropylene bags (50 kg each) per acre. A maximum of 200 jute bags (or 400 polypropylene bags) is issued to a farmer at a time.⁷ After the farmer returns filled bags, another consignment of *bardana* is issued, if needed. The farmer must pay a deposit of Rs. 134 per jute bag (or Rs. 38 per polypropylene bag) into a designated bank account. The bank prepares a call deposit receipt, which the farmer presents at the Centre to receive the *bardana*. Farmers recoup this deposit upon delivery of bagged wheat.

Wheat-surplus farmers usually try to dispose off their produce as quickly as possible since they need cash for next sowing and PFD's interventions create little incentive for investment in storage (Prikhodko and Zrilyi, 2013). PFD being the preferred buyer due to its high price, issuance of *bardana* becomes a locus of control and patronage. PFD issues *bardana* preferentially to those who are well-integrated into local political and social networks. Farmers keep local PFD officials in good humor throughout the year; the favour is returned at the harvest time. One farmer told us: "In our lives, local PFD officials matter almost as much as the local police officials and *patwaris* do. We cannot afford to annoy any of them." In years of bumper crop and low international price, PFD follows an unofficial go-slow policy in *bardana* issuance. For example, in 2017 and 2018, the Punjab government, unwilling to buy the entire stock at a procurement price that it found politically unfeasible to reduce, wriggled out of the situation by rationing *bardana* to a favoured small set of farmers. The rest were put in a slow-moving queue. Three of our farmer respondents whose names appeared on the PFD list could not get *bardana* in their first couple of visits, whereafter they sold their wheat to middlemen at the (lower) market price.

Farmers who do get the *bardana* can deliver wheat-filled bags only to their designated centre. They fill their bags, weigh them, stitch them, load them in a trolley and drive/escort them to the Centre. Not every farmer has the infrastructure to carry out these (seemingly simple) activities. The trolley has to be rented, and the rent has to be paid in cash. An ordinary trolley can carry 100 bags. Only two of our small farmer respondents had a marketable surplus exceeding 100 bags.

Once the trolley arrives at the Centre, it is weighed if the weighbridge is available, or a 10 percent sample is drawn and weighed to calculate the total weight. The sample is also assessed for quality. It is not uncommon for politically misaligned farmers to experience more rigorous quality testing. The In-charge Fountain House told us:

⁷ In 2018, the limit was set at maximum of 80 bags per farmer.

“In 2016, we had a good crop and our name was on the PFD list. So, I filled the bags and loaded the entire stock (about 200 maunds) onto trollies and sent these to the Centre. I am on good terms with the Centre In-charge. Still, he was unusually strict. He said our grain size was smaller than officially prescribed. When my staff argued with him, he asked them to approach the Deputy Director, PFD, which we did. The Deputy Director was a good man. When we told him, we are a charity and spend our entire income on mentally patients, he phoned the Centre In-charge to take it easy.”

Transportation to the Centre and unloading of bags is farmers’ responsibility. To help with these labour intensive and backbreaking activities, casual labour is available at the centre, which is paid by farmers at Rs. 9 per 100 kg bag. PFD reimburses these labour charges, though not farmers’ expense on filling and stitching of bags, loading and transportation to the centre.

While issuing *bardana*, PFD officials also advise farmers a preferred date for wheat delivery. The Centre populates its calendar in a manner that procurement operations are spread over the entire procurement season (mid-April to end-May). If everything goes by the schedule, the transaction is usually completed within 5-6 hours. If it is an unusually busy day, or if the farmer is ahead of or behind schedule, it may take longer. The farmer is not compensated for the wait.

PFD also shortchanges farmers by procuring a little extra wheat (without payment) to compensate for loss during storage and transportation. PFD’s total storage capacity is 2.19 million tons, and the rest of the procured wheat is stored in the open under plastic sheets. PFD officials estimate the loss accruing in covered and open storage as 2-2.5 percent and 5-6 percent respectively. Still, PFD operates under a zero-loss presumption (Prihodko and Zrilyi, 2013). One Centre In-charge commented: “our seniors know that there will be some loss during operations, yet they expect us to make up the loss somehow magically.” The staff readily shifts the burden onto the farmer. All farmers in our sample complained about being shortchanged. Fountain House, for example, had 200 maunds for sale to PFD in 2016. They were shortchanged one kg/maund, which translated into a loss of Rs. 6,500 for the entire consignment.

Upon delivery of wheat, PFD calculates farmer’s payment and pays cash for smaller consignments (< 50 bags) or issues a slip for larger consignments to be presented at the designated bank branch. At day’s end, PFD prepares a consolidated statement for the bank’s convenience. The bank pays the farmer upon receiving the consolidated statement and farmer’s payment slip, which means a farmer cannot receive his payment the same day but must visit the bank the next day or later.

Our farmer respondents who sold their wheat to PFD in 2017/2018 found this process bureaucratic and arduous. They had to visit the procurement Centre *at least* thrice – once to ascertain if their name was in the list, then for obtaining *bardana* and finally for delivering wheat. They had to visit the bank *at least* twice – once to pay call deposit for *bardana* and then to collect payment. This process took between 7-10 days. Throughout this process, they looked up to various PFD officials for approvals – for issuance of *bardana*, for quality clearance and for payment. They also faced a lot of uncertainty. Even when their names were on the list, they did not know if PFD would still be procuring when their turn came. PFD stops procuring upon meeting its target for the year.

Ostensibly, these practices are designed to regulate procurement but in practice they place a substantial administrative burden on the farmer. They comprise a mechanism to effectively exclude those very small farmers who are the *raison d'être* of this public intervention. Cash-strapped farmers who need quick payment to settle a debt, to buy inputs for the next crop, or to finance a marriage, sometimes prefer to sell to intermediaries for this very reason. One farmer respondent said:

I had only 15-20 days after the wheat harvest during which I had to plough my land to clear it of weeds, and to get it levelled. I don't have a tractor, so I rent it. But I can either rent one before everyone else needs them or after they are done. I can't possibly do the rounds to banks and food centres.

Further, small farmers are less likely to have the required surplus that would justify multiple visits to various offices, *in situ* labour of bag filling, weighing and stitching, and transportation. They are also less likely to be adequately networked to navigate their way through the official maze. In 2017, when the wheat campaign was being launched, the then Chief Minister publicly vowed that 'small farmers will be given preference in wheat procurement and the government would procure as much produce as they would bring to the procurement centres' (*Dawn* 5th April, 2017). This public pronouncement from the highest level notwithstanding, three of our small farmer respondents and the Fountain House failed to sell their wheat to PFD in that year. The bureaucratic procurement processes had effectively excluded them.

These bureaucratic procedures also create plenty of rent seeking and patronage opportunity for PFD officials. For example, in both 2017 and 2018, market prices were substantially lower (Rs. 1,100 – 1,250) than the PFD price (Rs. 1,300). In these years, while everyone wanted to sell to PFD, it had substantial carry forward stocks (Table 1) and was unwilling to procure beyond its target. In the absence of a neutral queuing system, this demand-supply gap enabled PFD officials in the field to procure wheat from their favoured farmers especially as the procurement season was ending. Although, none of our farmer respondents reported paying a bribe or using a reference (i.e. *sifarish*) themselves, almost everyone knew someone who had.

A key challenge, therefore, for PFD is to reduce considerable discretion exercised by these officials. Recent debates about bureaucratic accountability and policy implementation highlight the discriminatory use of discretion and how this shifts policy benefits away from marginalised groups (Epp, et al. 2014; Schram, et al. 2009). An important enabler for misuse of discretionary authority is that traditional modes of monitoring, such as field inspections, audits, etc., fail to monitor every interaction between frontline officials and citizens. Only when officials stop acting like what Bovens and Zouridis (2002) call 'tiny oligarchs' benefits of policy interventions reach their intended target groups. For PFD, one possibility is to use technology to improve monitoring of its frontline officials, say, by developing and following a neutral queuing system that brings greater certainty to the procurement process. Another is to simplify business processes and make them more transparent. Reduction of discretion at the frontlines is a key challenge for PFD if it wants micro-processes of procurement to include, rather than exclude tenants and small farmers.

5. THE REAL BENEFICIARY

PFD's wheat subsidy is built into the (usually) above-market procurement price and below-market issue price (i.e. the price at which it sells wheat to flour mills). There have

been years in which market price was higher than the official price, forcing PFD to ban inter-provincial movement to meet its procurement targets (e.g. 2014 (*The Nation*, 14th May, 2014). Generally, however, PFD sets a high procurement price to incentivise wheat cultivation, and issues wheat to flour mills at a discounted price (PFD Overview, n.d.).

PFD's subsidy regime has the following objectives: (1) reduce flour price volatility; (2) protect farmers from a market crash; and (3) provide subsidised wheat flour to consumers (PFD office records).⁸ PFD has been largely successful in reducing seasonal fluctuations in flour prices (Pieters and Swinnen, 2016, PBS, 2015: Table 219). Pieters & Swinnen (2016) show that annual fluctuations across the years have also been reduced. However, PFD has only partially succeeded in achieving the other two objectives (discussed shortly). Yet the cost to the government is substantial.

In 2017-18, the total cost of PFD's wheat operations was Rs. 34.43 billion (Table 2). Ironically, such huge cost—collectively called incidental charges—was mainly due to bank mark-up on money borrowed for procurement. 2017-18 was by no means exceptional; PFD's cost of wheat operations per unit was simply a continuation of a decade-long trend. Bank mark-up accounted for the largest share of the total incidental cost during the past decade: it was > 90 percent in 2010-13, declined thereafter, but was still quite high (around 70 percent) in 2017-18. The next two items were transportation and PFD salaries, which together accounted for a mere 9.7 percent of the total cost (Table 3).

Table 2

Total Cost of PFD's Wheat Operations

	08-09	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18
Procurement (million ton)	5.78	3.72	3.19	2.78	3.68	3.74	3.23	3.93	3.95	3.62
Total Cost (billion Rs.)	14.45	18.60	23.33	22.42	29.38	21.42	22.78	37.26	26.68	34.43
Cost per Ton (Rs.)	2,500	5,000	7,313	8,065	7,984	5,727	7,052	9,480	6,755	9,511

Source: Constructed from PFD data.

Table 3

Breakdown of Cost (Rs. per Metric Ton)

Component	08-09	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18
Gunny Bags	12	13	(413)	(116)	(681)	5	272	1,177	788	1,284
Delivery Expenses	70	75	75	75	75	75	75	75	90	75
Bank Commission	59	89	89	89	99	113	113	122	122	122
Taxes & Duties	12	16	16	12	11	12	14	15	11	15
Transportation Charges	885	871	98	338	556	502	602	602	497	769
Handling Charges	13	1	3	1	3	-	2	2	-	2.55
Godown Expenses	90	444	133	140	202	237	385	392	222	431
Storage & Unforeseen Expenses	-	-	145	-	123	-	-	-	-	-
PFD Charges	145	212	188	261	330	266	293	320	308	352
Interest (as % of Total Incidentals)	1,214 (49)	3,279 (66)	6,978 (95)	7,264 (90)	7,266 (91)	4,427 (77)	5,295 (75)	6,775 (72)	4,717 (70)	6,461 (68)
Total per Metric Ton	2,500	5,000	7,313	8,065	7,984	5,727	7,052	9,480	6,755	9,511
Total per 40 kg	100	200	293	323	319	229	282	379	270	380

Source: constructed from PFD data.

Note: These are actual costs which have not been adjusted for inflation.

⁸ Another objective is to increase domestic production to reduce imports and to protect consumers from international price fluctuations.

Overall, PFD's wheat operations have been progressively costing more. The per ton cost rose from Rs. 2,500 in 2008-09 to Rs. 9,511 in 2017-18. Though, this is largely due to an increase in the mark-up charges, the increase in other items is not insignificant altogether. For example, PFD charges have increased from Rs. 145 per ton to Rs. 352 per ton. Similarly, bank commission is now more than twice of what it used to be ten years ago. This increase is worrisome since it implies that procuring wheat is becoming increasingly expensive for PFD.

The unusually high mark-up charges in recent years warrant explanation. PFD has a credit line with three consortia of banks to be utilised during the procurement season. The amount is repaid during the year as wheat is sold to flour mills. Not wanting to pass on the full cost of operations to consumers, the Punjab government picks up part thereof as a subsidy. However, since 2008 the government has been clearing only a part of its liabilities due to financial constraints and the rest has remained outstanding each year (see Table 4). Except in 2010-11 and 2012-13, repayment has been less than new borrowing, thereby increasing the outstanding liability. Consequently, PFD has had to pay the mark-up not only on its current borrowing but also on the outstanding debt. On average, during the last five years, PFD has paid a whopping Rs. 19.6 billion each year as mark-up alone. *Prima facie*, the Punjab Government has no intention of paying back the outstanding dues, as it has budgeted only Rs. 10 billion each year since 2014 for PFD's wheat operations (PFD office records). Apparently, there is a tacit understanding that only mark-up will be paid each year—that too partly—and the outstanding liability will be allowed to stay—and mount.

Table 4

Borrowing and Mark-up (Million Rs.)

	Bank Borrowing			Repayment	Balance	Mark-up Paid
	Outstanding	New	Total			
2008-09	—	52,394	52,394	45,795	6,599	3,942
2009-10	6,599	143,176	149,775	47,376	102,400	19,028
2010-11	102,400	88,667	191,067	101,067	90,000	25,971
2011-12	90,000	76,025	166,025	56,127	109,898	23,180
2012-13	109,898	73,305	183,203	104,258	78,945	20,233
2013-14	78,945	110,540	189,485	97,267	92,218	16,272
2014-15	92,218	112,580	204,798	39,098	165,700	19,827
2015-16	165,700	105,345	271,045	75,231	195,814	18,492
2016-17	195,814	128,064	323,879	100,740	223,139	18,536
2017-18	223,139	128,846	351,985	22,942	329,043	21,465
2018-19	329,043	115,444	444,487	—	—	—

Source: constructed from PFD data.

By 2018, the outstanding debt had increased to a level that alarmed both PFD and the Punjab Finance Department. PFD office records show that both departments opposed large procurement in 2018. The latter even sent a summary to the Chief Minister in March, 2018 before the Punjab cabinet met to set a procurement target for the year. The summary noted that there was hardly any justification for large procurement in 2018

when a stock of 3.59 million ton was already available with PFD and Rs. 329 billion were owed to banks on account of previous wheat operations. The Finance Department suggested to keep new procurement strictly to whatever PFD needed (additional to existing stocks) for issuance to flour mills over the year and using the saving to retire the outstanding debt (PFD office records). But with the general elections due in a few months, the advice was a little untimely. The Chief Minister promptly overruled it. Instead, PFD was given a procurement target of four million ton for 2018 (Dawn, 15th April 2018).

For the consortia of banks, things could not have been better. Lending to the government is the least risky and has little transaction costs, if any. Banks get a competitive interest rate from a borrower who is happy to pay the mark-up year after year on a growing portfolio. Their principal amount is secure against sovereign guarantees. It is hard to escape the conclusion that banks have been a major beneficiary of PFD's wheat operations during the past decade.

6. ACTORS AND THEIR INTERESTS

According to PFD's pronouncements (e.g. *Dawn* 5th April 2017), its wheat operations are carried out for the benefit of small farmers. But the above discussion shows that the process is anything but friendly for small farmers. Are small farmers still benefitting, or are there other groups that are benefitting and are exerting enough policy influence to continue PFD's wheat operations despite not making sense otherwise? To answer this question, we will have to critically examine the interests of key actors, such as farmers, intermediaries, PFD, flour mills and consumers.

Farmers

An important reason for government intervention is to avoid a market crash in April-May when farmers are trying to dispose off their produce. However, farmers are not a homogenous group and they benefit variously from government intervention according to what Manig (1990) calls the 'societal modes of distribution.' Below we look at official statistics on farm size and wheat production to broadly comment on the benefit of PFD's wheat operations on various farmer categories.

Data on wheat production by farm size in Punjab (Tables 5 and 6) show that 13.27 percent of farms do not produce wheat at all. Of the 86.73 percent farms that produce wheat (comprising 63.71 percent of the total area of private farms), 90.45 percent are smaller than 12.5 acres, i.e. 78.45 percent of total farms producing wheat are small farms. These farms are only 40.50 percent of the total farm area and 63.56 percent of the area of farms reporting wheat.

Very small farms (<1 acre; 8.94 percent of all farms in Punjab) also do not sell wheat because they do not have a marketable surplus. Data on wheat production by farm size are not available, but assuming that they produce at the national average of 800 kg wheat per acre and consume wheat flour @ 140 kg per capita (Dorosh and Salam, 2008), and have an average household size of 6.3 (HIES, 2015-16), they produce less than their own consumption even when they allocate the entire area to wheat production. Together, non-wheat-producing farms and very small farms are 22.21 percent of all farms in Punjab.

Table 5

Wheat Area by Farm Size in Punjab

Farm Size (Acre)	Total Farms		Farms Reporting Wheat		Wheat Farms as % of Total Farms	
	No.	Area (acre)	No.	Area (acre)	No.	Area (acre)
< 1	729,981	306,511	469,412	152,333	8.94	0.52
1 – 5	2,617,507	6,205,234	2,363,425	4,600,168	45.02	15.69
5 – 12.5	1,412,603	10,478,386	1,285,613	7,078,499	24.49	24.14
12.5 – 25	359,408	5,812,909	316,642	3,518,254	6.03	12.00
>25	130,322	6,523,401	117,968	3,289,459	2.25	11.22
Total	5,249,821	29,326,441	4,553,060	18,683,533	86.73	63.71

Source: Constructed from Pakistan Agriculture Census 2010 (Tables 1.1 and 6.11).

Table 6

Wheat Producing Farms by Farm Size

	Farms in Punjab	
	No.	%
Total Farms	5,249,821	100
Farms do not Report Wheat	697,176	13.27
Farms Reporting Wheat		
< 1 Acre (too small to have a surplus)	469,412	8.94
1 – 12.15 Acre (mostly excluded by the arduous process)	3,649,028	69.51
> 12.5 Acre (have marketable surplus and networks)	434,610	8.28

Source: Constructed from Pakistan Agriculture Census 2010 (Tables 1.1 and 6.11).

Farms of 1-12.5 acres are 69.51 percent of total farms and represent 39.83 percent area. These farmers are likely to have a marketable surplus, but the PFD protocols discussed above effectively discourage most of them from selling their produce to the government. All eight farmers of <12.5 acres that we interviewed testified to this effect. Four of them in 2017 and five in 2018 sold their wheat to local traders. Other respondents, including PFD officials and grain merchants also observed that farmers in this category are more likely to sell in the market than to PFD.

The last category is farms larger than 12.5 acres. Our field work shows that these are PFD's real clients. These large farms represent only 8.28 percent of total farms but 23.22 percent of the area. These farmers produce marketable surplus and aspire to sell at official price should their local networks enable them to do so. Dorosh and Salam (2008) estimate that the top 10 percent wheat farmers in terms of sales account for 47 percent of total wheat sales and the top 20 percent account for 67 percent of total wheat sales (though not exclusively to PFD). Wheat sales (to PFD and elsewhere) are concentrated at the very top!

Our PFD respondents were of the view that these large farmers use their political clout to lobby with the government for an ever-increasing procurement price and large annual procurement targets. This is how PFD explained the Punjab government's refusal to accept their suggestion in 2018 for a gradual reduction of procurement targets (discussed above). Our field work shows that these farmers also ingratiate themselves with local PFD officials so that they get the *bardana* on a priority basis, their waiting

time is the least at the Centre as they are frequently allowed to jump the queue, and few objections are raised to the quality of their harvest. In the words of one large farmer respondent:

I am friends with the Deputy Director in PFD. He regularly visits me on various functions. The entire staff knows this. So, they give me *bardana* promptly and as per my requirement. Sometimes, I get *bardana* and pay the deposit later on. I never face a problem.

The relationship is not one-sided, though. In deficit years, i.e. when PFD is struggling to meet its procurement targets due to high demand in other provinces or abroad (e.g. in 2014 (*The Nation* 14th May 2014)), these farmers come to PFD's rescue. 'When everyone else is trying to make money, we honor our friendships,' said one farmer respondent. In these years, they help local PFD officials meet their targets and reap the returns in surplus years in the shape of preferential treatment. Such a symbiotic relationship is difficult to establish and maintain for small farmers simply because their marketable surplus is too small to be significant for PFD in deficit years.

In recent past, PFD has attempted to discourage preferred buying from large farmers through including a landownership ceiling in eligibility criteria (*Dawn* 15th April 2018) or restricting the number of bags a farmer can receive in one go. However, PFD officials were skeptical about the effectiveness of these measures because land records indicate ownership rather than the operational size of the farm. The latter is usually larger than the former because most landholdings are joint but are operated by a single family member. So, for example, if a landowner of 20 acres leaves behind five legal heirs with equal shares, the land record will indicate five farmers of four acre each, whereas on the ground the farm is more likely cultivated by one family members whereas others will seek employment elsewhere, if they can. In other words, on record, there are five small farmers but, on the ground, there is one large farmer and four persons with non-farm incomes. PFD officials opined that this dissonance between the *de jure* and the *de facto* undermined their efforts to restrict PFD buying to small farmers using land ownership as the criteria.

Estimating the benefit incidence of PFD's wheat operations on any group of farmers is difficult. There are large supply-demand variations from year to year. In years of high (anticipated) consumption, the market price may be very close to, or even higher than the official price. In May 2014, for example, our farmer respondents in Sheikhpura district sold their produce at the farm-gate to intermediaries at Rs. 1,190 per 40 kg as against the official price of Rs. 1,200 per 40 kg. In surplus years, however, the market price may be significantly lower than the official price. For example, the same farmer respondents sold their produce at Rs. 1,160 and Rs. 1,105 per 40 kg in May 2017 and April 2018 respectively as against a procurement price of Rs. 1,300. Further, the market price also varies according to the farm's location. A metric ton in a wheat surplus district is cheaper than a metric ton in a wheat deficit district, simply because the former has to be ultimately transported to high-consumption areas.

High procurement price benefits wheat-surplus farmers including those who do not sell to PFD. It positively impacts farmers' bargaining position in the market. A rising tide lifts all ships. The World Bank (2015) estimates that in the absence of any buying at all by the government, the market price may fall to Rs. 950 per 40 kg in a surplus year. The drop will be lower if the government does not withdraw from the market but procures smaller quantities.

This was the case in 2017 and 2018—both years of a bumper crop and large carry forward stocks in PFD granaries. PFD adopted an unofficial go-slow policy and rationed *bardana* to politically networked farmers. Consequently, the market price fell to Rs. 1,150 – 1,180 per 40 kg in various districts of Punjab in 2017 (compared to the official price of Rs. 1,300 per 40 kg). Losing hope in PFD procurement, seven of our farmer respondents sold to local traders at the going price (only four sold to PFD; remaining four used their wheat for own and extended family's consumption). This saved them (approximately) Rs. 20 per 40 kg for the cost of stitching, loading and transportation, as traders collected wheat from the farm-gate and brought their own bags. So, the net difference between a farmer selling to PFD and a farmer selling to a local trader in 2017 was approximately Rs. 100–130 per 40 kg. This was, then, the approximate benefit that accrued to those of our farmer respondents who sold wheat to PFD in 2017. This benefit was not insignificant, but it was confined to medium and large farmers. All four farmers in our sample who sold to PFD in 2017 had farms larger than 12.5 acres. The situation was very similar in 2018. The market price was even lower and only these same four farmers sold to PFD. Another way of putting this is that a system designed and implemented in the name of small farmers ended up excluding them from a benefit of Rs. 100–130 per 40 kg.

But this is not to suggest that small farmers or those medium farmers who did not sell to PFD did not benefit at all. Assuming WB estimates on price-drop in case of complete PFD disengagement were correct, farmers who sold to intermediaries also benefited between Rs. 200 – 230 per 40 kg.

Beoparis

Local traders (called *beoparis*) purchase at the farm gate and pay cash. *Beoparis* use their own *bardana*, and arrange for loading, transportation and unloading. They provide all these services for their profit at the intersection of market and procurement prices. The *beoparis* we interviewed reported selling small quantities of wheat to PFD in both 2017 and 2018. They completed all formalities on behalf of the farmers they purchased wheat from. To this extent, PFD records were erroneous since they showed purchase from farmers. On being asked if PFD officials ever noticed that they were not purchasing from farmers, one *beopari* replied:

They just care about the farmer's National Identity Card, and as long as we are able to provide that, there is no problem. But it has now become difficult since payment is also to be received at the bank.

Mostly, however, these *beoparis* stocked wheat in the harvest season and sold to retailers, bakers and flour mills subsequently as prices rose gradually.

Do *beoparis* make excessive profit? Anecdotal evidence from our fieldwork suggests that they do not. *Beoparis* provide a range of services and maintain a network of client farmers from whom they purchase wheat on annual basis. If the difference between their cost price and the official price was large in most years, a larger number of investors would enter the market and bring down profits.⁹ If it was insignificant, *beoparis* would

⁹ There are no significant barriers to entry. All that it takes to become a *beopari* is to have the capacity to buy wheat and to store it for a few months. Most *beoparis* operate small-scale. Large buyers are flour mills and Commission Agents (Arhti) in grain markets.

not invest their money and time in wheat operations. For example, as noted above the difference between market and procurement price was Rs. 110-130 per 40 kg in 2017. So, *beoparis* buying from the farmer saved in the range Rs. 90-110 (deducting Rs. 20 for stitching, etc.) per 40 kg. This is around 7 percent on their investment, which is neither insignificant nor excessive. In due course prices rise. The highest they went in 2017 and 2018 was Rs. 1,540 per 40 kg. *Beoparis* make money here as well depending upon their storage and transaction costs.

While small farmers benefit from a high procurement price, *beoparis* benefit from the arduous procurement process. If the process is simplified, a larger proportion of small farmers will prefer to present their produce directly at the Centre in a surplus year. On the other hand, any additional procedural requirement will persuade more of them to sell their stocks in the open market. To the extent the process is simplified, they will have a meaningful alternative and their bargaining power *vis a vis* the *beopari* will improve.

PFD

During our fieldwork, we came across several instances when rent seeking and patronage opportunities created by the bureaucratic process were exploited by unscrupulous staff. The loose accountability framework for government officials encourages system's manipulation for personal gain. Our farmer respondents testified, and PFD senior management concurred, that in surplus years, it was not uncommon for PFD officials to exploit their discretion in issuing *bardana*, quality control, weighing, etc. to strengthen their personal networks. These networks were subsequently deployed to gain choice postings, escape audits and avoid accountability. In the words of a senior PFD official: "We have some nuisance value, admittedly less than the police but arguably more than, say, the livestock department. People like being posted in PFD." Overall, these 'micro practices of corruption' (Krishnamurthy, 2012) place PFD officials as another beneficiary of the wheat operations in Punjab.

Clearly, PFD's interest lies in the continuation of large-scale procurement, involving the bureaucratic procedures. We found little evidence during our interaction with PFD officials of any soul-searching to re-engineer the business process, reduce transaction costs, improve quality of service, or reimagine the subsidy regime. Senior officers were, however, visibly perturbed at the ever-rising cost of debt servicing. PFD Secretary formally advised the Chief Minister Punjab in 2017 and again in 2018 to retire outstanding debt and reduce procurement targets (PFD office records). The advice was quietly ignored, as it required diversion of resources from priority projects and risked offending powerful middle/large farmer lobby.

Flour Mills

Flour mills¹⁰ profit from PFD's wheat operations in three ways. First, they receive their annual capacity-based wheat quota at a below-market price. As noted above, wheat prices rose to a peak of Rs. 1,540 per 40 kg in 2017-18, but flour mills received wheat from PFD at a price of Rs. 1,300 per 40 kg (Table 7). Second, PFD moves wheat at its

¹⁰There are 910 mills in Punjab. 80 percent are small-to-medium size.

Table 7

Subsidy per 40 kg (Rs.)

	Procurement Price (1)	Incidentals (2)	Cost Price (1+2=3)	Release Price (4)	Subsidy (3-4=5)
2008-09	625	100	725	545-900	180-(-175)
2009-10	950	200	1,150	975-1,000	175-150
2010-11	950	293	1,243	975-1,000	268-243
2011-12	950	323	1,273	1,000	273
2012-13	1,050	319	1,369	1,075-1,125	294-244
2013-14	1,200	229	1,429	1,330	99
2014-15	1,200	282	1,482	1,280-1,330	202-152
2015-16	1,300	379	1,679	1,300	379
2016-17	1,300	270	1,570	1,300	270
2017-18	1,300	380	1,680	1,300	380

Source: Constructed from PFD data.

expense from areas of procurement to areas of consumption. Wheat production is concentrated in Southern and Western Punjab, whereas consumption is spread across the province. On average around 0.7-0.8 million tons are moved annually to ensure that sufficient stocks are available in each district to meet allocated quota of mills (PFD office records). PFD bears the transportation cost (Rs. 60-70 per 40 kg). Third, mills sell flour at the standard price regardless of whether wheat is sourced from PFD over the year or bought cheaply the open market at the harvest time. These profits are substantial (Dorosh and Salam, 2008).

PFD sets the flour price periodically as a function of the issue price, transportation costs, grinding costs and millers' margin. The issue price may remain the same or vary through the year. There are substantial rent opportunities for mills that receive subsidised wheat from the government and sell flour at the market price (Dorosh and Salam, 2008). Both millers we interviewed reported grinding cheap wheat purchased from the farmer during May-September but selling flour at the (higher) official price. From October to April, they grind PFD's subsidised wheat and sell at the official price. As long as PFD bears the cost of storage, wastage and transportation, they have little incentive to buy larger quantities during harvest season.

Consumers: PFD issues wheat to flour mills at prices lower than its cost (Table 7). In 2016 and 2017, for example, PFD issued wheat to flour mills on the procurement price. This means the entire cost of operations, storage, transportation, etc. was picked up by PFD as a subsidy to consumers. In other years, the general trend has been to pick up part of the incidentals in the hope that this subsidy would be passed on to consumers.

This is not to suggest that the cost absorbed by PFD is equivalent to consumer subsidy, which would be the difference between the consumer price of wheat flour with and without PFD intervention. We make no effort to estimate that. All that is said above (and in Table 7 below) is that PFD picked up the entire cost of its wheat operations (Tables 2 and 3) as a subsidy to consumers in 2015-17. If PFD improves its efficiency or reduces its borrowing costs, it can provide wheat to flour mills at the same price as in 2017, but at a smaller cost to itself.

Consumers are a diverse group and various sub-groups are affected differently. Wheat-deficit rural households and urban consumers are directly affected by a change in flour prices. According to official data (PBS, 2015: Table 7), wheat products account for 12.84 percent of total household expenditure in Punjab. Bottom two quintiles spend 20.65 percent and 17.73 percent respectively of their income on wheat flour (PBS, 2015). These quintiles are vulnerable to the adverse effects of price increases.

Will the price of wheat flour rise by the amount of subsidy if PFD stops wheat operations? Certainly not. Instead, flour price may decrease due to two reasons. First, in the absence of public intervention, the wheat price will fall at the harvest time and traders will stock up at these low market prices. Second, the private sector's wastage and storage/transportation costs will be lower than PFD incidentals (Aamir, 2017). Thus, mills' net cost will be lower than it is in the current interventionist regime.

What will be the net effect if PFD procures wheat but without having to pay interest on the piled-up outstanding debt? Data in Table 4 show that current borrowing in 2017 accounted for only 36.60 percent of the total PFD debt. If PFD had to pay mark-up only on its borrowing for current operations, its incidentals would be correspondingly less. Adjusting the mark-up as shown in Table 3 accordingly gives us a figure of Rs. 205.19 per 40 kg for PFD incidentals, were there no outstanding debt. If PFD procures wheat at the official price and if it pays mark-up only on new borrowing, its cost price of wheat at the time of issuance to flour mills will be Rs. $(1,300 + 205.19 =) 1,505.19$ per 40 kg. If PFD provides no subsidy at the issuance of wheat, cost of flour will increase by Rs. 205.19 per 40 kg; if it provides subsidy at the current level (viz. Rs. 379 per 40 kg), cost of flour will decrease by Rs. $(205.19 - 379 =) 173.81$ per 40 kg, *ceteris paribus*. In other words, consumers are paying Rs. 173.81 extra per 40 kg for Punjab government's failure to retire its outstanding debt.

7. CONCLUSIONS AND POLICY RECOMMENDATIONS

We have documented above how banks, small/middle/large farmers, *beoparis*, flour mills and PFD officials benefit from Punjab's wheat subsidy regime. By looking at the micro-processes of procurement and by dissecting numbers, we make a case for reimagining the wheat subsidy regime so that a greater incidence of benefit falls upon small farmers and consumers in whose name the subsidy is maintained. In 2017-18, the total cost of PFD's wheat operations was Rs. 34.43 billion. This is not an insignificant amount for a province whose total budget in that year was Rs. 1,970 billion (Punjab, 2018). The similar cost incurred year after year is like a small, non-fatal wound that bleeds the body imperceptibly and steadily. This 'friendly fire' consumes valuable resources, excludes small farmers and deprives the consumers of a potential benefit.

The question immediately arises: when all this is known to the government, why does it not reform the subsidy regime? PFD leadership admitted that government's reluctance emanated from fears that reform would result in low wheat prices in the harvest season and high and volatile flour prices in the rest of the year, both of which may lead to violent unrest. These fears are not totally unfounded. Global evidence suggests that subsidy reforms may trigger a series of events that are potentially destabilising, especially when latent political discontent exists (Gutner, 2002; Hopkins, 1988). Amid (2006) has cited several cases where subsidy reforms provoked popular

unrest. Therefore, gradual reform will be a safer strategy for the government, starting with improved targeting, open sharing of ideas, and awareness of the general public on the reasons behind actions being taken.

Government's reform options depend on its objectives. If the objective is to address the short-term problem of the government's mounting fiscal burden, it is imperative to repay the high-cost, short-term debt to commercial banks. Procurement targets better aligned with PFD's releases to flour mills will enable PFD to clear each year's liabilities within that year. Clearly, PFD should not allow its outstanding liabilities to mount again. As seen above, this serves neither the government nor its target groups. If the objective is to develop a more efficient wheat market, the government should gradually reduce its footprint on the wheat landscape. Instead, it should strategically intervene to regulate the interaction amongst farmers, *beoparis*, mills and consumers. If the objective is to address the structural issue of small-farmer exclusion, the government should reform the business process. These objectives are not mutually exclusive.

The way forward emerges from our discussion above. First, the government should improve the evidence base for price setting and strengthen PFD's analytical capacity so that procurement prices and targets reflect market requirements and PFD's strategic objectives and fiscal space. In general, the procurement price should be better aligned with the international price. This price should be kept flexible so that it can move both ways in response to market conditions. If the government liberalises international trade and restricts regulation of import/export to exceptional circumstances, international prices will provide the hedge against inflationary pressure on domestic prices. However, as noted by Pieters and Swinnen (2016) and Alonso and Swinnen (2016), linking domestic prices to international prices will transmit the latter's volatility to the former. This will defeat an important purpose of PFD's wheat procurement regime. To avoid such an outcome, the government should continue to intervene in the market strategically. Second, the government should initiate measures to offset the adverse effects of price liberalisation in the short run. We propose the creation of a Wheat Support Fund, to which should go savings from reduced PFD costs. This Fund should be used exclusively for improving wheat-related rural infrastructure, and for supporting the development of wheat storage by farmers in particular and the private sector more generally. Currently, even large farmers and traders have little incentive to invest in storage facilities as they cannot compete with massively subsidised PFD operations. But a reduced PFD role will create the space for farmers/traders to develop their own storage facilities. If PFD staggers its procurement over the May-October period and provides a premium for later procurement, this will incentivise middle/large farmers to hold on to their wheat for longer, rather than dispose it off at once in the harvest season. Small farmers will benefit indirectly, as there will be correspondingly less wheat for the market to clear in the harvest season. Staggering procurement will also reduce PFD's cost of storage and mark-up charges to the extent that PFD procures smaller quantities in April-May. The premium paid for later purchases should be smaller than PFD's cost in case of earlier purchases. Third, PFD's business process should be improved to make the procurement process less problematic especially for small farmers. PFD should consider accepting wheat in farmers' *bardana* or at least restricting its *bardana* issuance to farmers as per small farmers' production capacity. Finally, as earlier advocated by Dorosh and Salam (2008),

PFD should issue wheat to flour mills on market rates through open auction, rather than on capacity-based quotas and arbitrary prices. This will help PFD recover some of the costs of its wheat operations. This will also increase mills' incentive to procure larger quantities from farmers at the harvest time and store these under their own arrangement for later use. To this extent, procurement, transportation and storage functions will transfer from PFD to flour mills.

A reduced PFD footprint may benefit consumers through lower wheat flour prices at least in surplus years when domestic production is plenty and export opportunities are scanty. Consumers will benefit from increased wheat purchases by flour mills and *beoparis* at harvest time. To protect low-income consumers against price volatility and higher flour prices in deficit years, the government may consider direct cash provision. It should be possible to target such cash provision to the bottom two quintiles through the existing databases of poor households (e.g. the one maintained by the Benazir Income Support Programme). A key challenge will be to regularly update the database and enhance bureaucratic capacity to identify households most in need of cheap flour. There is plenty of evidence from across the globe suggesting that targeted subsidies may help the poor at lower costs (Alderman, et al. 2018; Gentilini, 2016; Lofgren and El-Said, 2001; Margolies and Hoddinott, 2015; Subbarao, et al. 1997). However, care must be taken that the cost of alternative programs is not higher than the generalised subsidy that they will replace.

REFERENCES

- Aamir, S. (2017). Policy reforms and analysis of wheat procurement system in Punjab, Pakistan. Munich Personal RePEc Archive. (Working Paper No. 81562).
- Ahmad, M. and Farooq, Umar (2010). The state of food security in Pakistan: Future challenges and coping strategies. *The Pakistan Development Review*, 49(4).
- Alderman, H. (1988). The twilight of flour rationing in Pakistan. *Food Policy*, August, 245–256.
- Alderman, H., Gentilini, Ugo, and Yemtsov, Ruslan (2018). *The 1.5 billion people question: Food, vouchers, or cash transfers?* Washington, DC: The World Bank.
- Alonso, E. and Swinnen, Johan (2016). Who are the producers and consumers? Value chains and food policy effects in the wheat sector in Pakistan. *Food Policy*, 61, 40–58.
- Amid, J. (2007). The dilemma of cheap food and self-sufficiency: The case of wheat in Iran. *Food Policy*, 32, 537–552.
- Aziz, S. (1979). Abolishing hunger: The complex reality of food. *Third World Quarterly*, 1(4), 17–27.
- Babbie, E. (2012). *The practice of social research* (13th Edition). California: Wadsworth Publishing.
- Bovens, M. and Zouidis, S. (2002). From street-level to system-level bureaucracies: How information and communication technology is transforming administrative discretion and constitutional control. *Public Administration Review*, 62(2), 174–184.
- Cummings, R., Rashid, S., and Gulati, A. (2006). Grain price stabilisation experience in Asia: What have we learned? *Food Policy*, 31(4), 302–312.
- Dawn (2017). <https://www.dawn.com/news/1324980/punjab-to-procure-4m-ton-wheat>

- Dawn (2018). <https://www.dawn.com/news/1401695>
- Dorosh, P. and Salam, Abdul (2008). Wheat markets and price stabilisation in Pakistan: An analysis of policy options. *The Pakistan Development Review*, 71–87.
- Epp, C. R., Marnard-Moody, S., and Haider-Markel, D. P. (2014). *Pulled over: How police stops define race and citizenship*. Chicago: University of Chicago Press.
- Gentilini, U. (2016). Revisiting the ‘cash versus food’ debate: New evidence for an old puzzle? *World Bank Research Observer*, 31(1), 135–67.
- Gutner, T. (2002). The political economy of food subsidy reform: The case of Egypt. *Food Policy*, 27(5-6), 455–476.
- Hopkins, F. (1988). The politics of adjustment: The African case. *Food Policy*, 13(1), 47–55.
- Islam, Yassir and Garrett, James (1997). *IFPRI and the abolition of the wheat flour ration shops in Pakistan*. Washington, DC: International Food Policy Research Institute.
- Krishnamurthy, M. (2012). States of wheat: The changing dynamics of Public procurement in Madhya Pradesh. *Economic and Political Weekly*, 47(52), 72–83.
- Lofgren, H. and El-Said, M. (2001). Food subsidies in Egypt: Reform options, distribution and welfare. *Food Policy*, 26, 65–83.
- Manig, W. (1990). Formal and informal credit markets for agricultural development in developing countries—The example of Pakistan. *Journal of Rural Studies*, 6(2), 209–215.
- Margolies, A. and Hoddinott, John (2015). Costing alternative transfer modalities. *Journal of Development Effectiveness*, 7(1), 1–16.
- O’Connor, J. (1979). *The fiscal crisis of the state*. NJ: Transaction Publishers.
- Offe, C. (1984). *Contradictions of the welfare state*. Cambridge MA: MIT press.
- Pakistan, Government of (2016). *Pakistan economic survey 2015-16*. Ministry of Finance, Government of Pakistan.
- Punjab, Government of (2018). *Budget in brief*. Punjab Finance Department.
- PBS (2015). *Punjab development statistics 2015-16*. Lahore: Punjab Bureau of Statistics.
- PFD Market Stabilisation (n.d.). https://food.punjab.gov.pk/market_stabilization. Accessed 11/04/2018.
- PFD Overview (n.d.). <https://food.punjab.gov.pk/overview>. Accessed 12/03/2018.
- Pieters, H. and Swinnen, Johan (2016). Trading-off volatility and distortions? *Food Policy*, 61, 27–39.
- Prikhodko, Dmitry and Zrilyi, Oleksandre (2013). *Review of the wheat sector and grain storage issues*. Rome: Food and Agriculture Organisation.
- Schram, S. F., Soss, J. Fording, R. C., and House, L. (2009). Deciding to discipline: Race, choice, and punishment at the frontline of welfare reform. *American Sociological Review*, 74(3), 398–422.
- Scott, J. and Hernandez, Citlalli (2018). From food subsidies to targeted transfer in Mexico. In Alderman, H., Gentilini, Ugo, & Yemtsov, Ruslan (eds.) *The 1.5 billion people question: Food, vouchers, or cash transfers?* Washington, DC: World Bank.
- Sen, A. (2005). *The argumentative Indian: Writings on Indian history, culture and identity*. New York: Farrar, Straus and Giroux.

- Subbarao, K., Bannerjee, A., Braithwaite, J., Carvalhe, S., Ezemenari, J., Graham, C., and Thompson, A. (1997). *Safety net programme and poverty reduction: Lessons from cross-country experience*. Washington, DC: World Bank.
- The Nation* (2014). <https://nation.com.pk/14-May-2014/punjab-bans-inter-provincial-movement-of-wheat>
- World Bank (2010). *Food price increase in South Asia: National responses and regional dimensions*. Washington, DC: World Bank.
- World Bank (2015). Strengthening markets for agriculture and rural transformation (SMART – PforR): Policy note for secretary agriculture. (Unpublished).

Remittances and Healthcare Expenditures: Evidence from Pakistan

KARIM KHAN, MUHAMMAD JEHANGIR KHAN, and ABID HUSSAIN

This study, examines the effect of remittances on healthcare expenditure in Pakistan by utilising the Pakistan Social and Living standards Measurement (PSLM) survey. The total healthcare expenditure is classified into two categories, i.e. expenditure on medicines and expenditure on clinical services. The study analyses these categories in case of both rural and urban areas of the country. Such data is generally characterised by selection bias; therefore, we employ Propensity Score Matching (PSM) instead of the commonly used econometric techniques. Findings of the study indicate that remittances enhance spending on both the clinical services and medicines. This result is robust across the urban and rural areas of Pakistan. The comparison between the clinical services and medicines shows that the impact is higher on clinical services as compared to the impact on medicines. This suggests that remittances help to improve the preventive nature of health outcomes in Pakistan.

JEL Classification: D10; I21; O15

Keywords: Remittances, Healthcare Expenditure, Medical Expenditure, Clinical Expenditure, Propensity Score Matching

1. INTRODUCTION

Over the time, labour market has been globalised and, as a result, migration for employment has increased. Accordingly, remittances have emerged as important source of income, especially in developing countries, which play an important role in enhancing consumption, private investment, education and health expenditure etc. (Ruiz-Arranz and Giuliano, 2005; Mundaca, 2009; Ahmed, et al. 2018). According to the World Migration Report (2020), the world's migrant population is estimated at around 272 million in 2019, which comprises 3.5 percent of the world's population. Likewise, during the last decade, we have experienced 2.3 percent growth in world's migration which constitute as twofold of that of the previous decade, implying an increasing trend in world migration. Consequently, the inflow of remittances to the receiving countries has also been increasing. In 2018, the inflow of global remittances was \$689 billion, out of which \$529 billion was directed to the low and middle-income countries. This large amount of transfers has clear implications for the overall economic development of receiving

Karim Khan <karim.khan@pide.org.pk> is Associate Professor, Pakistan Institute of Development Economics, Islamabad. Muhammad Jehangir Khan <jehangir@pide.org.pk> is Assistant Professor, Pakistan Institute of Development Economics, Islamabad. Abid Hussain <abidhussain_12@pide.edu.pk> is Data Coder O/O Revenue Office (P) Bahawalnagar MEPCO (Multan Electric Power Company).

countries; however; it largely depends on the household contexts, community circumstances, and the way decisions are made (Adams and Page, 2005; Kalaj, 2015). Pakistan is no exception in this regard. In 2019, the migrants from Pakistan were around 6.3 million which roughly constitute as 2.9 percent of the country's population. In 2019, according to the World Bank (2019), the receipts of remittances to Pakistan was approximately \$22 billion which places Pakistan in the top ten remittance-receiving countries. Given the importance of remittances for Pakistan, a considerable research has been conducted to explore the macroeconomic impact of remittances.¹ For instance, most of the existing research has been conducted in the context of rising remittances and its relationship to consumption expenditure, poverty, economic growth etc. To our knowledge, we are the first to study the impact of remittances on health care expenditure in Pakistan from micro perspective.

With regard to the impact of remittances on healthcare expenditure; there are a few studies available in the existing literature which claim that remittances enhance healthcare expenditure (Amuedo-Dorantes and Pozo, 2011; Jorge, 2008; Clement, 2011; Kalaj, 2015). In general, remittances affect the health of households in several ways. First, remittances relax the credit constraint of households which, in turn, results in the increase in healthcare expenditure of the households (Lopez, et al., 2007). Indirectly, migration increases the awareness and knowledge about the health standards and, thus, enhances the direct effect of remittances on health (Lindstrom and Munoz-Franco, 2006; Hildebrand and McKenzie, 2005). Health, as a key element of human capital and future productivity, has significant implications for economic prosperity and poverty alleviation (Grossman, 1972). In general, individual's demand for health is directly related with their incomes. In this study, we intend to examine whether this direct relationship holds for remittances? Arif (2004) studies the impact of remittances on health-related activities in Pakistan. Abbas, et al. (2014) make similar attempt by looking at their implications for healthcare expenditure. However, these studies are characterised by smaller sample size and are only restricted to rural areas. In addition, these studies have selection-bias problem due to the omission of observable characteristics of households.

Keeping in view these problems, we study the effect of remittance on households' healthcare expenditure at micro-level. We decompose this impact into two categories. First, we study the implications of remittances for households' clinical expenditure which comprises annual fees paid to specialists, doctors, Hakeem/Midwives, laboratory test, hospital charges, etc. Second, we analyse the effect of remittances on the expenditure of households on medicine which includes the annual expenses on the purchase of medical apparatus, vitamins, equipment or supplies and medicines etc. In order to do this analysis, we use Pakistan Social and Living Standards Measurement (PSLM) (2011-12) survey by employing the approach of Propensity Score Matching (PSM). In order to see the effects of regions, we separately analyse the urban and rural areas in terms of healthcare expenditure. Rest of the study is categorised into four sections. Section 2 discusses the relevant literature. Methodology, the construction of variables and Data have been discussed in Section 3. Section 4 shows the main results along with the discussions on those results while Section 5 concludes the study.

¹ See, for instance, Bilquees and Hamid (1981), Arif (2004), Abbas, et al. (2014) etc.

2. REVIEW OF LITERATURE

In general, health is considered as an asset which can be produced. So, the production of health is an investment which counter-balance the consumption of capital (Zweifel, et al. 2009). The spending in health is rewarded by the reduction in time spent in bad health. Thus, the demand for medical care is a derived demand which can be perceived in similar pattern as the demand for other commodities (Grossman, 1972).² Households spend on the health services only if such spending is relatively advantageous compared to the spending on other commodities. Alternatively, households spend on health services if the job hours saved due to good health are more beneficial as compared to the amount spent on healthcare expenditure (Zweifel, et al. 2009). Thus, as investment, cost-benefit analysis is done for health just like other investment projects. In addition to the pricing and earning strategies, there are many other demographic characteristics which determine health demand like the household structure; family, economic and social contexts; structure of the market for workers and wages; government institutional structure etc. (Ke, et al. 2011). Parental education is also important among the other determinants of health demand. Toor and Butt (2005) find that the parents with higher level of education are likely to take more care of their children, resulting in better health services for their kids.³ Groot, et al. (2006) assert that parental education indirectly affect health demand through its effect on children's education.⁴ With regard to regional effects, Hotchkiss, et al. (1998), while matching the budgetary position, find that in urban areas the healthcare expenditures are relatively small as compared to the rural areas.

There is considerable research which focuses on the impact of migration or remittances on the provision of healthcare or healthcare spending.⁵ Lindstrom and Munoz-Franco (2006) find that remittances enhance healthcare expenditure in Guatemala.⁶ Alternatively, remittances relax the credit constraint and, accordingly, remittances receiver households are likely to spend more on health care. Amuedo-Dorantes and Pozo (2011) got similar results in case of Mexico.⁷ Likewise, in case of Mexico, Valero-Gil (2009), finds that households which have no access to employment medical services spend 11.3 percent of the increase in remittances on health care as compared to 8 percent for those which have approach to employment medical services.⁸

²The Grossmann's model is a benchmark model of the demand for healthcare expenditures. According to the model, individuals demand high quantity of health services if the labour wages and health capital are high. Moreover, there is indirect relationship between the demand for health services and their prices. Likewise, the education level has a negative relationship and age has a positive relationship with quantity demanded of health services.

³For instance, due to parental care, the incidence and length of illness is significantly decreasing. Likewise, parental education increases the trust on modern medical treatments as compared to outdated medicines. See, also, Odubunmi and Abidogun (2013), Groot, et al. (2006) for the details.

⁴Alternatively, parental education indirectly affects the demand for health by taking into account the implications of better health for educational success of the children.

⁵See, for instance, Lindstrom and Munoz-Franco (2006), Amuedo-Dorantes and Pozo (2011), Jorge (2008), Drabo and Ebeke (2011), Clement (2011) etc.

⁶In particular, the effect in the rural areas of Guatemala is larger.

⁷The decomposed analysis of Amuedo-Dorantes and Pozo (2011) shows that the healthcare expenditure is less sensitive to remittances in the lower-income households. However, households which lack healthcare coverage show higher response to remittances. See also Frank, et al. (2009) for the case of Mexico.

⁸These referred to various benefits for permanent employees provided after their regularisation. It is based on medical coverage for employee and her/his dependents.

In a similar study, Jorge (2008) concludes that, in the absence of workers' insurance policies for the clinical services, almost 10 percent of the increase in migrants' remittances are allocated to expenditure on healthcare services. In a slightly different study, Drabo and Ebeke (2011) find that foreign aid, public health spending, and remittances play key role in the access to healthcare services in developing countries. In this regard, remittances lead to a sectoral glide to the private sector from the public sector for the richest and intermediate income classes. In sharp contrast to the above studies, Ponce, et al. (2011) concludes that remittances enhance the total consumption and healthcare expenditure; however, they have no significant effect on child health. As we can see, most of the existing literature asserts that remittances enhance healthcare expenditure and, thereby, have significant impact on the health status of the members of the households living in the home country.

Similar attempts have been made in Pakistan in order to analyse the effect of remittances and migration on healthcare expenditure. For instance, Arif (2004) analyses the impact of migration on health-related activities by controlling for other demographic and socio-economic characteristics of households.⁹ The study finds that the infant mortality rate is higher in households which have migrants; and this effect is larger for female kids. With regard to remittances, Abbas, et al. (2014) find that remittances enhance total expenditures on food, education and health in Punjab, Pakistan. In contrast to these two studies, Bilqees and Hamid (1981) find that the remittances result in the reduction of poverty in general; however, its impact on health is negligible as there is significant difference between the migrants and non-migrants households in terms of healthcare expenditure. In this study, we use a larger sample and incorporate both the rural and urban areas. Also, we take significant care for sample selection bias and endogeneity issues by employing the approach of Propensity Score Matching (PSM).

3. DATA AND METHODOLOGY

This section throws light on the characteristics of data along with the methodological framework for analysis. In particular, this section explains the quasi-experimental approach. Finally, the section explains variables, including both the dependent and explanatory variables.

3.1. Description of Data

We use Pakistan Social and Living standards Measurement (PSLM) survey, 2011-12, in order to draw data for our analysis.¹⁰ PSLM includes data on a variety of aspects including income, education, employment, health and migration at provincial and national levels. Thus, it is one of the most suitable nationally representative surveys in order to examine the effect of remittances on expenditures on healthcare in Pakistan. There are separate questionnaires for males and females. Females' questionnaire gathers information related to health, maternity history, pregnancy (for married females of 15-49

⁹The analysis is based on the data from Pakistan Socio-economic Survey (PSES) for 2001. The study focus on migration instead of the amount of remittances received from the migrants.

¹⁰Pakistan Bureau of Statistics (PBS) conduct PSLM.

years of age), family planning, reproductive health etc. besides demographic and socio-economic information of the females. Likewise, males' questionnaire gathers information on household expenditure, socio-economic and demographic characteristics of the households etc. PSLM 2011-12 covers a sample of about 15807 households which are distributed over 1158 Primary Sampling Units (PSUs) (585 urban and 573 rural). According to the survey; 861 are remittances-receiving households which constitutes as 5.4 percent of the total households. 358 (41.6 percent) of these households are from the urban areas while 503 (58.4 percent) are from the rural areas. During 2011-12, sickness or injury is observed in 7 percent of the total population. Around 96 percent of the total injured or sick households had experienced some health consultations, out of which 22 percent benefited from public facilities (i.e., hospitals, Rural Health Centers (RHCs), dispensaries, and Basic Health Units (BHUs) while 71 percent availed private hospitals/dispensaries.

A brief comparison of remittances-receiver and non-receiver households in terms of socio-economic characteristics is shown in Table 1. The household size for both the remittances-receiver and non-receiver households is similar in urban areas; however, it is significantly different in case of rural areas and the whole sample. Remittance-receiver households have a higher proportion of female heads (43 percent) than non-receiver households (8 percent). This proportion largely remains the same across the rural and urban areas of the country. With respect to age of the households' heads, heads of remittance-receiver households are older than those of non-receiver households. The proportion of uneducated heads are higher in remittance-receiving households as compared to non-receiving households, but this difference is significant only in case of the whole sample and that of the urban areas. Remittance-receiving households have higher healthcare expenditure than non-receiving households for the whole sample as well as for the rural and urban areas. The total health expenditure at the disaggregate level, in terms of medicine expenditure and clinical expenditure, also shows significant differences between non-receiving and receiving households. Finally, remittance-receiver households are higher in Punjab (51 percent) followed by KPK (45 percent).

3.2. Method for Measuring Impacts

There are two approaches to measure impacts, i.e. experimental and non-experimental. In our case, due to the unavailability of data for experimental approach, we opt for quasi-experimental approach which is the most used method of non-experimental approach.¹¹ In quasi-experimental approach, the quasi-experiment is likely to produce an associated control group by questioning: "what the treatment group may have done if the treatment was not done?" (Armendariz and Morduch, 2010). The randomisation procedure in quasi-experimental approach is done in such a way that all the observed and unobserved characteristics except the treatment are similar for both the control and treatment groups (Rosenbaum and Rubin, 1983; Bryson, et al. 2002). There are three different methods in this approach: Before-After Difference Estimator (BADE); Matching Estimator; and Difference in Difference Estimator (DID). We avail matching

¹¹In fact, in impact-evaluation, the non-experimental programs are relatively easy and cheap to implement (Smith and Todd, 2005).

Table 1
Summary Statistics of Remittances-receiver and Non-receiver Households.

Variables	Whole Sample			Rural Sample			Urban Sample		
	Remit.	Non-Remit.	S.E. Diff.	Remit.	Non-Remit.	S.E. Diff.	Remit.	Non-Remit.	S.E. Diff.
Household Size	7.14	6.72	0.12(0.00)***	7.69	6.82	0.16(0.00)***	6.37	6.577	0.17(0.23)
Head Age	48.71	46.20	0.48(0.00)***	48.55	45.46	0.64(0.00)***	48.94	47.192	0.71(0.01)**
Head Gender (Male=1)	0.57	0.92	0.01(0.00)***	0.59	0.92	0.01(0.00)***	0.54	0.93	0.01(0.00)***
Uneducated	0.47	0.42	0.01(0.01)**	0.54	0.52	0.02(0.26)	0.36	0.29	0.02(0.00)***
Grade 1	0.14	0.15	0.01(0.42)	0.16	0.17	0.02(0.61)	0.12	0.14	0.02(0.48)
Grade 2	0.10	0.11	0.01(0.67)	0.12	0.10	0.01(0.41)	0.09	0.12	0.02(0.11)
Grade 3	0.16	0.15	0.01(0.66)	0.13	0.12	0.02(0.81)	0.21	0.20	0.02(0.63)
Grade 4	0.12	0.16	0.01(0.00)***	0.06	0.09	0.01(0.01)**	0.22	0.26	0.02(0.07)**
Healthcare Exp.	14487.4	7639.3	531(0.00)***	14065.9	7441.3	683.2(0.00)***	15079.7	7904.8	841.2(0.00)***
Medical Exp.	4235.6	3243.5	309(0.00)***	3942.1	3070.1	397.1(0.03)**	4647.9	3475.9	489.5(0.02)**
Clinical Exp.	10251.8	4395.9	383(0.00)***	10123.7	4371.2	461.7(0.00)***	10431.7	4428.9	649.9(0.00)***
Punjab	0.51	0.43	0.02(0.00)***	0.44	0.44	0.02(0.90)	0.60	0.43	0.03(0.00)***
KPK	0.45	0.19	0.01(0.00)***	0.53	0.20	0.02(0.00)***	0.32	0.18	0.02(0.00)***
Sindh	0.03	0.28	0.02(0.00)***	0.001	0.27	0.02(0.00)***	0.06	0.29	0.02(0.00)***
Balochistan	0.02	0.09	0.01(0.00)***	0.02	0.09	0.01(0.00)***	0.02	0.10	0.02(0.00)***
No. Observation	861	14946	15807	503	8561	9064	358	6385	6743

Note: The figures in parenthesis shows the corresponding probabilities of significance. ***, **, * shows significance at 1 percent; 5 percent; and 10 percent, respectively. Exp., Remit., Non-Remit. show expenditure, Remittances-receiver households, Remittances Non-Receiver households, respectively. S.E. Diff. shows standard errors differences.

estimator due to its convenience.¹² According to Dehejia and Wahba (2002), Propensity Score Matching (PSM) is more convenient (with lower bias) if the data satisfies three necessary conditions. First, the sample for both the treatment and control groups should be drawn from the same geographical location. Second, the data should be collected through the same questionnaire. Third, the data set should include a large set of variables related to the treatment and consequence variables. The data set that we use in this study satisfy all of these conditions; therefore, PSM is more suitable for our analysis.

In order to define the treatment, we first analyse the key factors that determine the receipt of remittances by consulting to multivariate analysis. This analysis comprises the binary logistic regression model:

$$Rem_i = \alpha_0 + \alpha_1 I_i + \alpha_2 HH_i + \alpha_3 Rg_i + \varepsilon_i \quad \dots \quad \dots \quad \dots \quad \dots \quad (1)$$

In Equation 1, the dependent variable Rem_i shows that whether the household i is remittances-receiver?¹³ I_i is the vector which comprises the individual characteristics of household head like age, education, and sex. HH_i shows the household characteristics like size of the household. Rg_i is a vector representing all other controlled factors like regions and provinces. Rem_i is dichotomous in nature, therefore, we consult to binary Probit regression model. Onwards, in order to estimate the effect of remittances, the difference in the outcomes among the controlled and target groups is measured, that is,

$$ATT == E(Y | D = 1, X_i) - E(Y | D = 0, X_i) \quad \dots \quad \dots \quad \dots \quad \dots \quad (2)$$

In Equation 2, ATT shows Estimated Average Treatment-on-Treated effect. Y is the outcome which is the health care expenditure in this study. D is the dummy which takes a value of 1 if the household is getting remittances; otherwise, it takes the value of 0. In order to avoid this possibility of Overt Bias, we control for the observable characteristics, (X_i) in estimating the models (Lee, 2005).¹⁴ Mosley (1997) identify that there might exist the likelihood of unseen bias between the target group and the control group. However, randomised selection in design-based studies rule out this possibility. Alternatively, randomisation eliminates the hidden bias by cancelling out the unobservable features of both the control and the target group. Propensity Score Matching (PSM) assures the similarity between the treatment and comparison groups in terms of the observed characteristics (Becker and Ichino, 2002; Dehejia and Wahba, 2002; Ravallion, 2003).¹⁵ For instance, based on their propensity scores, the observed variables between the comparison group and the treatment group are balanced. This

¹² PSM, based on observed features, firstly measure the propensity score for both the remittances-receiver non-receiver households. Onwards, it compares the average outcome of remittances-receiver with that of the matched remittances non-receiver. The control group usually comprises those households among all the non-receiver households which are similar to remittances receiver households in other characteristics. Alternatively, the control group would have the same result as compared to what the remittances receivers would have if they had received the remittances. This assumption is known as the Conditional Independence Assumption (CIA) (Rosenbaum and Rubin, 1983).

¹³ Rem_i is dichotomous in nature with which has two outcomes. 1 denotes the receiving household while 0 denotes the non-receiver.

¹⁴ Over Bias occurs if the observable features are not identical.

¹⁵ According to Dehejia and Wahba (2002), the matching selects those non-participants who have identical features to those of the participants. Propensity scores are actually the predicted probabilities of receiving the treatment which, in our, study is the receipts of remittances.

particular mechanism is justified by the fact that if a variable affects only the participation but not the outcome; then it is not necessary to control for the differences in this variable in the treatment group versus the control group. Likewise, if a variable does not affect the treatment but instead, only affects the outcome; then it is not necessary to control for that variable because of the fact that the consequence will not meaningfully be changed in the treatment versus control groups. Finally, the variables that neither affect the result nor the treatment are also relatively irrelevant. Thus, the characteristics that affect both the treatment and outcome are necessary for the matching process which needs to be incorporated in the model for deriving the propensity scores.¹⁶

3.3. Description of Variables

In this subsection, we define the variables that are used in this study. Firstly, we define the treatment and outcome variables. Onwards, we discuss the control variables.

3.3.1. Treatment and Outcome Variables

The receipt of remittances is a treatment variable in this study. In PSLM, there are detailed questions with regard to domestic as well as foreign remittances received by the households. In this study, we focus only on foreign remittances in order to see its impact on healthcare expenditure. The outcome variable in this study is households' healthcare expenditure in Pakistan which comprises medical expenditure and clinical expenditure. In PSLM, medical expenditures are regarded as annual purchase of medicines and vitamins, medical apparatus, supplies or equipment and other health related expenditures. Likewise, clinical expenditures include the annual fees paid to specialists, hakeems/midwives and doctors outside the hospital. Additionally, it also incorporates hospital charges and expenses on laboratory tests etc. By combining the medical and clinical expenditures, we get the total health expenses of the households.

3.3.2. Control Variables

There are three sets of explanatory variables. First, we include the individual level characteristics of the households' head like household heads' gender, age, education and, age square. Second, as household characteristic, we use household size. Finally, we incorporate regional characteristics in terms of dummies for the provinces. Gender, in this study, has been incorporated to control for the gender of the household head. Education is the most important variable among the indicators of human capital as far as the effect of remittances on healthcare expenditures is concerned. It is a general perception that educated individuals are likely to spend more fractions of remittances on healthcare expenditures compared to the non-educated individuals. In PSLM, the education of respondent has been asked as the highest class achieved. We categorised them into five main groupings, i.e. illiterate, primary, middle, secondary and tertiary. We also use household size in list of control variables. It is expected that larger households will spend less proportion of remittances on healthcare expenditures as compared to smaller households. In this study, we carry out separate analysis for rural and urban regions, so that to observe any regional differences in healthcare expenditures of remittance-

¹⁶ See Caliendo and Kopeinig (2008) for the details of practical guidelines.

receiving and non-receiving households. This variable is characterised as rural and urban areas whereas urban serves as the corresponding reference group.

4. EMPIRICAL RESULTS

This section presents the empirical results. First, we discuss the results for the overall sample. Onwards, we discuss the results for the urban and rural areas, respectively.

4.1. Impact of Remittances on Healthcare Expenditure in Whole Sample

We estimate the propensity score of remittances receipt through Probit regression. As is stated earlier, the receipt of remittances is the treatment which takes the value of 1 if foreign remittances are received by the household; otherwise, it takes a value of 0. In this specification, we include household head's age, gender and education level as individual level characteristics. Likewise, on household level, it incorporates household size. Besides, dummy variables for region and provinces are taken in order to measure the regional and provincial variation in remittances. Before conducting the PSM, we need two conditions that need to be met in order to calculate the *Average Treatment on the Treated (ATT)* effect on the basis of the propensity scores of households (Rosenbaum and Rubin, 1983). First, given the propensity scores, we need to balance the pre-treatment characteristics for the control and treatment groups.¹⁷ Second, we need the condition of un-confoundedness which states that given a set of observable correlates 'X' that are not affected by treatment; the potential outcomes 'Y' are independent of treatment assignment 'T'. After finding propensity scores through Probit regression, we use Radius Matching and Kernel Matching in order to measure the *Average Treatment on the Treated (ATT)*.¹⁸

Table 2 shows the results on the determinants of remittances. We incorporate only those variables in the estimation for which the conditions of balancing and un-confoundedness are satisfied. As we can see, the probability of remittances increases with the age of household head and household size. Higher household size enhances the probability of members of the households to be migrants. Likewise, remittances are higher in those households which are headed by females. Region does not play any significant role in the receipt of remittances. Education of all categories enhances the probability of remittances which is a standard result in the literature on the returns to education. The cross-province comparison shows that remittances are higher in Punjab and Khyber-Pakhtunkhwa (KPK) as compared to Balochistan while remittances in Sindh are comparatively lower. This finding is justified by the fact that we have higher migrant networks in Punjab and Khyber-Pakhtunkhwa (KPK).

¹⁷See Figure A1, Tables A2 and A3 in appendix A for details.

¹⁸PSM approach tries to capture the effects of different observed covariates X on participation in a single propensity score or index. Then, outcomes of participating and nonparticipating households with similar propensity scores are compared to obtain the program effect. Households for which no match is found are dropped because no basis exists for comparison. Radius Matching and Kernel Matching are the two different methods of PSM. These are the matching algorithms (scales) that are used to match beneficiaries with non-beneficiaries. These scales identify the comparison group by checking for balance in the characteristics of the treatment and comparison groups; and estimating the program effect. In the radius matching, a maximum propensity score radius is established, and all non-beneficiaries within the given radius of a beneficiary are matched to that beneficiary. Whereas, in case of the Kernel matching for each treated subject, a weighted average of the outcome of all non-beneficiaries is derived. The weights are based on the distance of the non-beneficiaries' propensity scores to that of the treated subjects, with the highest weight given to those with scores closest to the treated unit.

Table 2

Probit Model for Propensity Scores to find the Effect of Remittances (Whole Sample)

Variable	Coefficient	St. Error
Region (Urban=1)	0.020	0.801
Head Age	0.017***	0.003
Head Gender (Male=1)	-2.352***	0.089
Household size	0.091***	0.000
Education of the Head (Illiterate as Reference)		
Grade 1-5	0.388***	0.116
Grade 6-8	0.464***	0.132
Grade 9-10	0.626***	0.116
Grade 11 and above	0.468***	0.127
Province (Baluchistan as Reference)		
Punjab	1.459***	0.256
Sindh	-0.686**	0.318
Khyber Pakhtunkhwa	2.036***	0.255
Constant	-4.058***	0.297
LR ² χ^2	1274.58	
Prob> χ	0.000	
Observation	15807	

Note: *, **, *** Shows significance at 10 percent; 5 percent; 1 percent, respectively. Among 15807 households, 859 households get the remittances and 14837 households don't get remittances.

Given the results in Table 2, the results on Average Treatment on Treated (ATT) effect are shown in Table 3. We can see that the ATT effect on the total healthcare expenditure is significantly positive. For instance, the coefficient from the Radius method is 0.499 which means that receivers of remittances spend 49.9 percent more on overall health care expenditure compared to the non-receivers. Likewise, in case of Kernel method, the result shows that receivers of remittances spend 43.7 percent more on overall healthcare expenditures compared to the non-receivers. The estimated effect is almost similar across both of the matching methods. Furthermore, the treatment units and the number of controls are similar in both cases, which implies that the sample size in both the cases remain the same. Given the disaggregation, the ATT effect of remittances on medicine expenditure shows that remittances receivers spend 24.6 and 22 percent more on medicines compared to the non-receivers in Radius and Kernel methods, respectively. Likewise, in case of clinical expenditure, the ATT effect of remittances shows that remittances receivers spend 61 and 49.9 percent more as compared to non-receivers in Radius and Kernel matching, respectively. Over all, the impact on clinical expenditure is higher than that of the medicine expenditure. This finding shows the income effect which implies that, at lower levels of incomes, households refer to informal health care providers or quack doctors.¹⁹ Alternatively, with the receipts of remittances, households shift from informal healthcare providers to the clinical services for their health care.

¹⁹ There are significant fraction of Informal healthcare providers (IPs) in the health system of developing nations (Sudhinaset, et al. 2013).

Table 3
ATT Effect of Remittances on Healthcare, Medicine and Clinical Expenditures
 (Whole Sample)

	Healthcare Expenditure		Medicine Expenditure		Clinical Expenditure	
	Radius Method	Kernel Method	Radius Method	Kernel Method	Radius Method	Kernel Method
ATT	0.499***	0.437***	0.246***	0.220***	0.610***	0.499***
No. Treated	859	859	599	599	840	840
No. Control	14837	14837	11827	11827	14166	14166
T-value	14.80	12.21	14.80	4.21	14.51	11.15

Note: *, **, *** shows significance at 10 percent; 5 percent; 1 percent, respectively. Bootstrapped standard errors with 10,000 repetitions.

4.2. Impact of Remittances on Healthcare Expenditure in Urban Areas

In order to measure the propensity score of remittances receipt in urban areas, we include age, age-squared, education level and gender of the household head as individual level characteristics. Likewise, on household level, it incorporates household size. In order to compare provincial level variations in the receipt of remittances, we also incorporate provincial dummies. The conditions of balancing and un-confoundedness are met to measure the *Average Treatment on the Treated* (ATT) effect based on the propensity score of households.²⁰ After calculating propensity scores through Probit regression model, the *Average Treatment on the Treated* (ATT) effect is calculated while using the Kernel Matching and Radius Matching. Table 4 shows the results on the determinants of remittances in this case. Again, we include only those variables for the properties of balancing and un-confoundedness are satisfied. Remittances in urban areas are higher for those households which are headed by females. Likewise, households whose heads' education level is high school or above receive higher remittances in the urban areas. The results also show that remittances are increasing household size. Even, in the urban areas, Punjab and Khyber-Pakhtunkhwa (KPK) is performing better in terms of remittances receipt as compared to Balochistan while Sindh has lower incidence of remittances receipts.

In order to estimate the effect of remittances on healthcare expenditure in case of urban areas, we estimate the ATT which are shown in Table 5. In case of total healthcare expenditure, the coefficient of ATT through Radius matching is 0.407 which indicates that receivers of the remittances are spending 40.7 percent more as compared to the non-receivers in the urban areas of Pakistan. Likewise, Kernel matching shows that remittances-receivers are spending 43.5 percent more as compared to non-receivers. The estimated effect is almost similar across both the matching methods. Furthermore, the treatment units and the number of controls are similar in both cases which implies that the sample size in both cases remains the same. In case of medicine expenditures, remittances-receivers spend 23 and 32.5 percent more in urban areas as compared to non-receivers in Radius and Kernel matching, respectively. Likewise, remittances-receivers spend 50.1 and 44.3 percent more on clinical expenditure in Radius and Kernel matching, respectively. Thus, in urban areas, the impact of remittances is higher on clinical expenditure as compared to medicine expenditure.

²⁰ See Figure B1, Tables B2 and B3 in Appendix B for details.

Table 4

*Probit Model for Calculating the Propensity Scores to find the Effect of Remittances
(Urban Households)*

Variable	Coefficient	St. Error
Head Gender (Male=1)	-2.510***	0.018
Head Age	0.010**	0.005
Household Size	0.051***	0.018
Education of Head (Illiterate as Reference)		
Grade 1-5	0.174	0.197
Grade 6-8	0.089	0.219
Grade 9-10	0.508***	0.170
Grade 11 and above	0.360**	0.166
Province (Balochistan as Reference)		
Punjab	1.376***	0.372
Sindh	-0.413	0.426
Khyber Pakhtunkhwa	1.668***	0.376
Constant	-3.057***	0.202
LR ² χ^2	527.37	
Prob> χ	0.000	
Observation	6743	

Note: *, **, *** Shows significance at 10 percent; 5 percent; 1 percent, respectively. Among 6743 households, 358 households are getting remittances while the remaining 6385 households are not getting remittances.

Table 5

*ATT Effect of Remittances on Health Care, Medicine and Clinical Expenditure
(Urban Areas)*

	Healthcare Expenditure		Medicine Expenditure		Clinical Expenditure	
	Radius	Kernel	Radius	Kernel	Radius	Kernel
	Method	Method	Method	Method	Method	Method
ATT	0.407***	0.435***	0.230***	0.325***	0.501***	0.443***
No. Treated	357	357	258	258	349	349
No. Control	6312	6312	5160	5160	6040	6040
T-value	7.47	7.46	3.04	4.03	7.66	6.28

Note: *, **, *** shows significance at 10 percent; 5 percent; 1 percent, respectively. Bootstrapped standard errors with 10,000 repetitions.

4.3. Impact of Remittances on Healthcare Expenditure in Rural Areas

This section reports results for the rural areas of the country. Again, in the first step, we estimate the propensity score of remittances receipt which include age, age-squared, gender, and education levels of the household head as individual level characteristics. On household level, we incorporate household size. To capture provincial variations in the receipt of remittances, we control for provincial dummies. Again, even, in this case, the conditions of balancing and un-confoundedness are met.²¹ The ATT effects are estimated for those households having the same observable characteristics

²¹ See Figure C1 and Tables C2 and C3 in Appendix C for details.

while using Radius and Kernel matching. Table 6 shows the results of the corresponding Probit regression. The results show that remittances are higher in those households which are headed by females. The coefficient on household size is positive and significant, depicting that the probability of remittances receipts increases with the increase in household size. The likelihood to receive remittances increases in the age of the household head. The estimate of household head education is positive and significant for almost all education levels except grade 11 and above.

Table 6

Probit Model for Calculating the Propensity Scores to Find the Effect of Remittances (Rural Households)

Variable	Coefficient	St. Error
Household Size	0.108***	0.013
Head Age	0.020***	0.004
Head Gender (Male=1)	-2.203***	0.121
Education of Head (Illiterate as Reference)		
Grade 1-5	0.481***	0.144
Grade 6-8	0.645***	0.166
Grade 9-10	0.642***	0.160
Grade 11 and above	0.410	0.216
Province (Balochistan as Reference)		
Punjab	1.494***	0.353
Sindh	-1.498***	0.564
Khyber Pakhtunkhwa	2.263***	0.350
Constant	-4.560***	0.399
LR ² χ^2	792.59	
Prob> χ	0.000	
Observation	9064	

Note: *, **, *** Shows significance at 10 percent; 5 percent; 1 percent, respectively. Among 9064 households, 503 households are getting the remittances while 8561 households are not getting remittances.

In order to see the impact of remittances on healthcare expenditure, we estimated the corresponding ATT effects which are shown in Table 7. The results show that the ATT effect on total healthcare expenditure is positive in rural areas of the country. The coefficient of ATT through Radius method is 0.565 which shows that remittances-receiver households spend 56.5 percent more on total healthcare expenditure as compared to non-receiver households. Likewise, Kernel matching shows that remittances-receiver households spend 39.5 percent more on total health care expenditure as compared to non-receiver households. The number of controls and treatment units are also same in both cases. In the case of medicine expenditure, the ATT from Radius matching shows that remittances-receivers spend 25.9 percent more on medicine expenditures compared to the non-receivers. The estimates of ATT effect from Kernel method shows that remittances-receivers spend 10.2 percent more on medicine expenditures compared to the non-receivers. Similar is the case with clinical expenditure where remittances-receivers spend 68.6 and 51.1 percent more than non-receivers in Radius and Kernel matching, respectively. The cross-region comparison reveals that remittances receiver households in rural areas spend more on clinical expenditure as compared to remittances-receiver households in the urban areas.

Table 7

*ATT Effect of Remittances on Health Care, Medicine and
Clinical Expenditure in Rural Areas*

	Health Care Expenditure		Medicine Expenditure		Clinical Expenditure	
	Radius Method	Kernel Method	Radius Method	Kernel Method	Radius Method	Kernel Method
ATT	0.565***	0.395***	0.259***	0.102***	0.687***	0.511***
No. Treated	502	502	341	341	491	491
No. Control	8525	8525	6667	6667	8126	8126
T-value	13.26	8.71	3.96	4.40	12.57	8.76

Note: *, **, *** Shows significance at 10 percent; 5 percent; 1 percent, respectively. Bootstrapped standard errors with 10,000 repetitions.

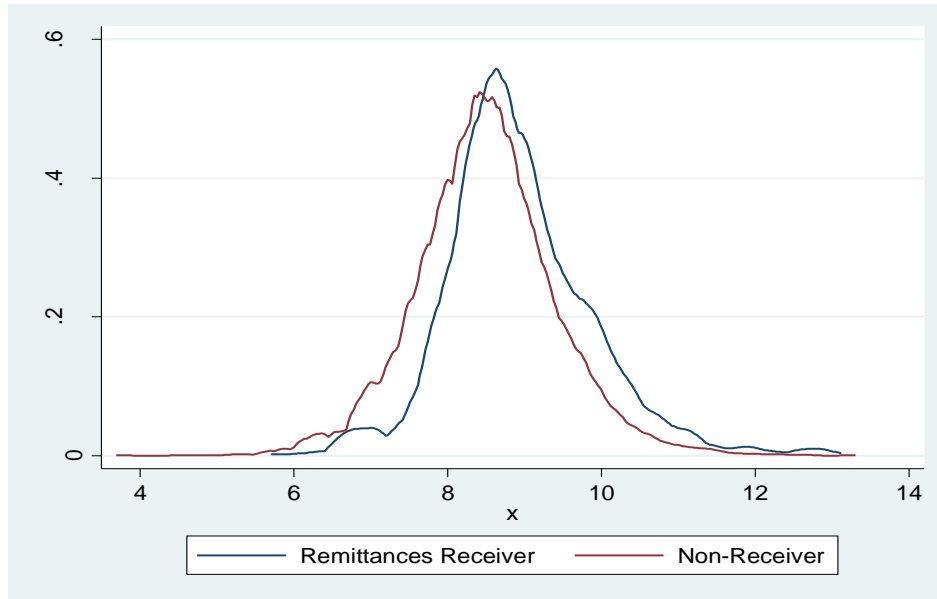
5. CONCLUSION

This study is motivated by recent surge in research on the effect of migration or remittances in developing countries. The inflow of remittances relaxes the credit constraints of households which, in turn, enhances spending on health care. In this study, specifically, we analyse the impact of remittance on household healthcare expenditure in Pakistan by using PSLM (2011-12). PSLM (2011-12) is one of the largest national representative survey which covers most of the socio-economic characteristics of households. This, in other words, implies that we cover a larger sample as compared to the existing studies. We classify total healthcare expenditure into medical expenditure and clinical expenditure. Firstly, we do the analysis for the whole sample for all the three categories, i.e. total healthcare expenditure, medical expenditure, and clinical expenditure. Onwards, we decompose our sample into urban and rural areas and do separate analysis for both the regions. In order to control for the problems of endogeneity and sample selection bias, we use the Propensity Scores Matching (PSM) instead of traditional econometric techniques.

We find that health expenditure, including both medical expenditure and clinical expenditure, are higher for the remittances-receivers as compared to the non-receivers. This positive relationship is statistically robust across the urban and rural areas of Pakistan. Moreover, for remittances-receiver households, the clinical expenditures are higher as compared to the medical expenditures. It is an indication of the fact that clinical services like clinical consultancies including fees of the doctors and specialists, laboratory tests, hospital charges etc. are expensive in Pakistan. Thus, remittances help in covering the expenditures on such services which illustrates improvement in the health status due to the precautionary nature of the health outcomes. In order to have better health outcomes in the country, our analysis suggests two guidelines. First, certainly, the government should encourage and, properly, channelise the inflow of remittances. Second, clinical services should be adequately regulated both for quality and user fees. However, certainly, more studies are needed to exactly identify the channels through which the effects of remittances translate into health outcomes. This would help us in providing clear policy guidelines in this regard.

APPENDIX A

Overlapping and Balancing Assumptions for Whole Sample

Fig. A1. Propensity Scores for Remittance Receiver and Non-receiver in Whole Sample

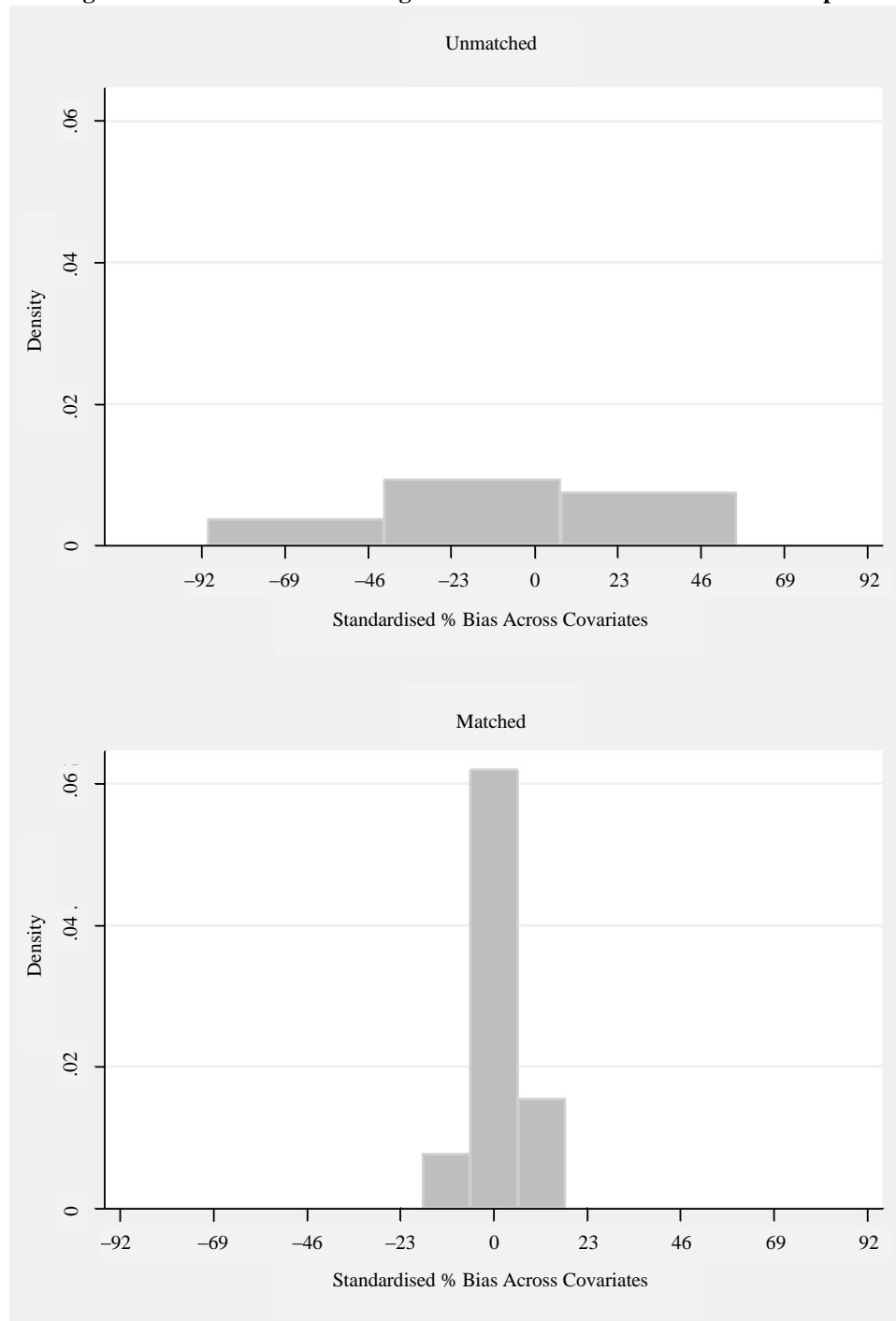
Note: Figure A1 shows the overlap between the distributions of the propensity scores of remittances receivers and non-receivers. PSM includes only those observations which lies in the overlapping areas while the remaining are dropped.

Table A1

Covariates between Treated and Non-treated Units of Remittances Receiver and Non-receiver (Whole Sample)

Variable	Mean Treated	Mean Untreated	St. Difference
Region(Urban=1)	0.51	0.43	0.175
Household Size	6.44	6.74	-0.080
Head Age	47.86	46.37	0.104
Head Gender	0.88	0.90	-0.073
Education of Head (Illiterate as Reference)			
Grade 1-5	0.12	0.16	-0.096
Grade 6-8	0.09	0.11	-0.046
Grade 9-10	0.15	0.15	-0.017
Grade 11 and above	0.16	0.16	-0.008
Punjab	0.47	0.44	0.068
Sindh	0.23	0.27	-0.100
Khyber Pakhtunkhwa	0.23	0.21	-0.282

Note: The second important assumption of PSM is to check the balance of covariates between treated and non-treated households. If covariates are not balanced between treated and non-treated households, the result will be upward or downward biased. In table A1, we can see that almost all of the covariates show balancing.

Fig. A2. Standardised Percentage Bias across Covariates in Whole Sample

Note: Figure A2 show the percentage reduction in bias after balancing the covariates between treated and non-treated households. Standardised percent bias is shrinking to about zero.

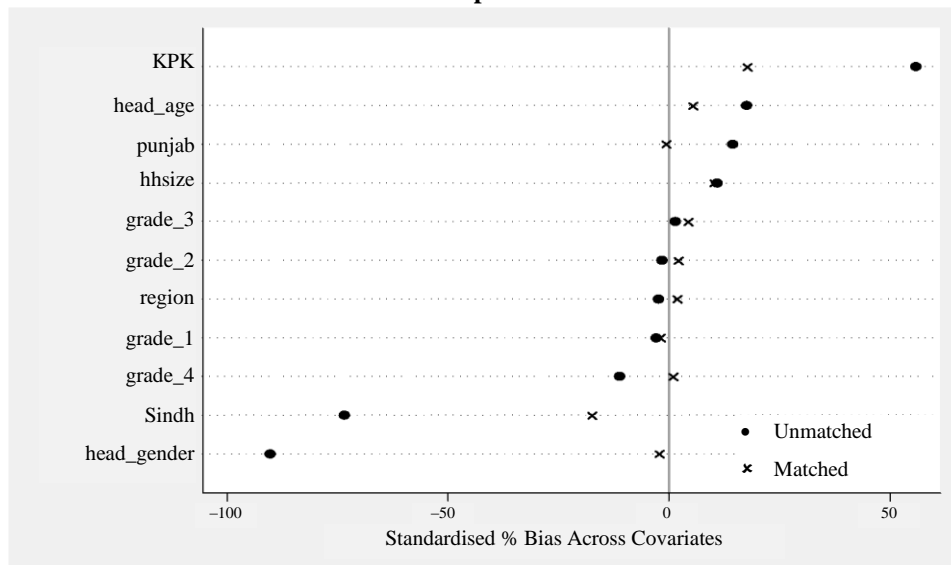
Table A2

Biasness between Unmatched and Matched Variables (Whole Sample)

Variable	Unmatched/Matched U/M	Mean		% Bias	Reduce Bias
		Treated	Control		
Head Gender (Male=1)	U	0.567	0.925	-90.2	97.2
	M	0.567	0.576	-2.2	
Household Size	U	7.137	6.719	11.0	7.3
	M	7.137	6.749	10.2	
Region(Urban=1)	U	0.416	0.427	-2.3	18.9
	M	0.416	0.407	1.9	
Head Age	U	48.714	46.198	17.6	69.1
	M	48.714	47.938	5.4	
Education of Head (Illiterate as Reference)					
Grade 1-5	U	0.144	0.154	-2.9	30.0
	M	0.144	0.151	-2.0	
Grade 6-8	U	0.105	0.109	-1.6	-35.8
	M	0.105	0.098	2.1	
Grade 9-10	U	0.159	0.154	1.5	-190.4
	M	0.159	0.143	4.4	
Grade 11 and above	U	0.124	0.163	-11.1	91.8
	M	0.124	0.121	0.9	
Province (Balochistan as Reference)					
Punjab	U	0.505	0.433	14.5	95.2
	M	0.505	0.509	-0.7	
Sindh	U	0.030	0.279	-73.3	76.1
	M	0.030	0.090	-17.5	
Khyber Pakhtunkhwa	U	0.445	0.194	56.0	68.3
	M	0.445	0.365	17.8	

Note: Table A2 shows the biasedness between mean of treated and non-treated households before and after the matching. The percentage reduction in bias between treated and no-treated households after matching is shown in the last column.

Fig. A3. Standardise Difference between Matched and Unmatched Variables in Whole Sample

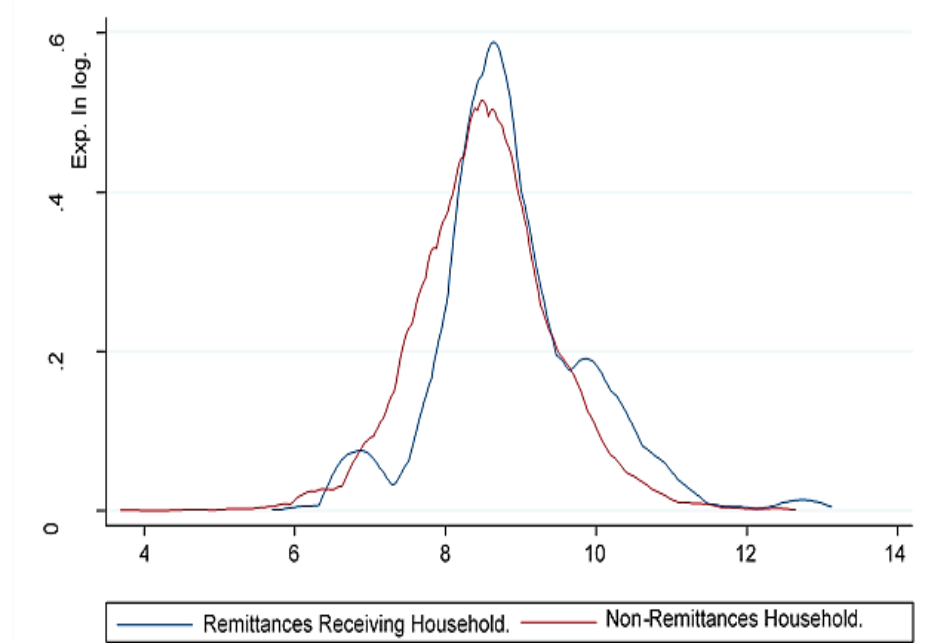


Note: Black dot represent the unmatched pair and cross dot represent matching pair. After using balancing technique of matching the bias is reduced.

APPENDIX B

Overlapping and Balancing Assumptions for Urban Areas

Fig. B1. Propensity of Scores for Remittance Receiver and Non-receiver in Urban Areas



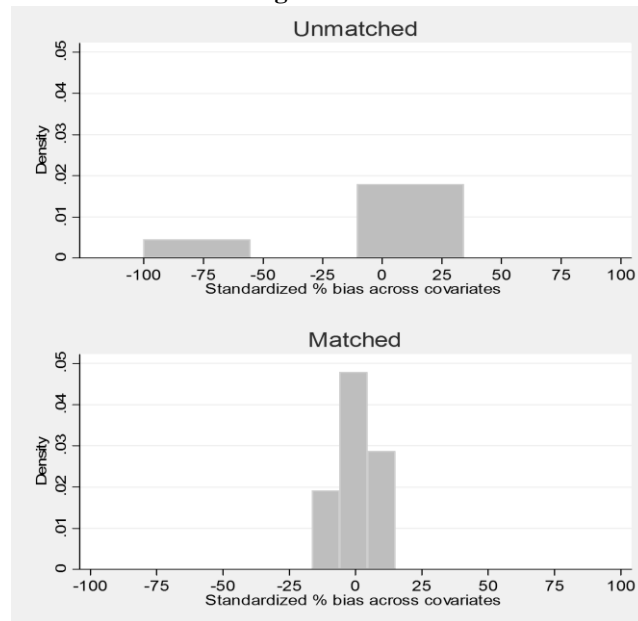
Note: Figure B1 displays that the overlap assumption is satisfied between non-treated and treated.

Table B1

Covariates between Treated and Non-treated Units of Remittances Receiver and Non-receiver (Urban Areas)

Variable	Mean Treated	Mean Untreated	St. Difference
Head Gender (Male=1)	0.89	0.91	-0.044
Household Size	6.15	6.57	-0.119
Head Age	50.33	47.34	0.216
Education of Head (Illiterate as Reference)			
Grade 1-5	0.13	0.14	-0.015
Grade 6-8	0.09	0.12	-0.097
Grade 9-10	0.18	0.20	-0.040
Grade 11 and above	0.22	0.26	-0.095
Punjab	0.50	0.44	0.122
Sindh	0.25	0.28	-0.092
Khyber Pakhtunkhwa	0.19	0.19	0.007

Note: In Table B1, we compare the variables between treated and non-treated households. The results show that head's gender of mean treated and untreated is almost the same. The variable household size's mean of untreated is greater than that of treated. Education of the head is in negative standardise difference in all grades and the provinces difference is same and positive.

Fig. B2. Standardised Percentage Bias across Covariates in Urban Areas

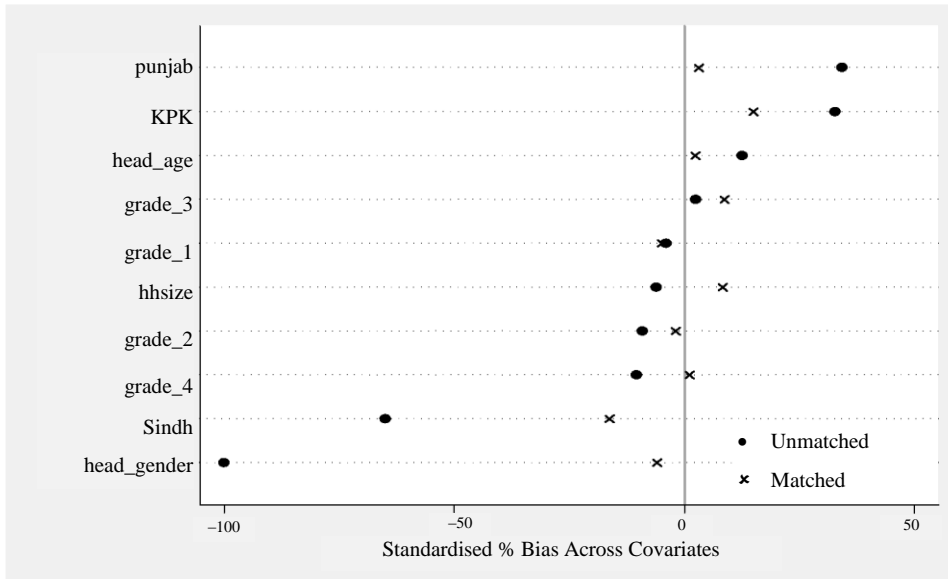
Note: Figure B2 clearly shows that bias in unmatched is greater as compared to bias in matched.

Table B2*Biasness between Unmatched and Matched Variables (Urban Sample)*

Variable	Unmatched/Matched U/M	Mean		%Bias	Reduce Bias
		Treated	Control		
Head Gender (Male=1)	U	0.539	0.936	-100	94.2
	M	0.539	0.563	-6.0	
Household Size	U	6.366	6.577	-6.0	-38.2
	M	6.366	6.074	8.3	
Head Age	U	48.941	47.193	12.7	81.8
	M	48.941	48.623	2.3	
Education of Head (Illiterate as References)					
Grade 1-5	U	0.123	0.136	-3.9	-25.4
	M	0.123	0.139	-4.9	
Grade 6-8	U	0.089	0.117	-9.0	78.7
	M	0.089	0.095	-1.9	
Grade 9-10	U	0.207	0.196	2.6	-235.7
	M	0.207	0.172	8.6	
Grade 11 and above	U	0.218	0.262	-10.3	88.8
	M	0.218	0.213	1.1	
Province (Balochistan as Reference)					
Punjab	U	0.598	0.428	34.4	90.9
	M	0.598	0.582	3.1	
Sindh	U	0.059	0.294	-64.9	74.9
	M	0.059	0.118	-16.3	
Khyber Pakhtunkhwa	U	0.321	0.180	32.9	54.3
	M	0.321	0.257	15	

Note: Table B2 show the biasedness between mean of treat and non-treated households before and after the matching in urban areas. The percentage reduction in bias between treated and no-treated households after matching is shown in the last column.

Fig. B3. Standardise Difference between Unmatched and Matched Variables (Urban Areas)

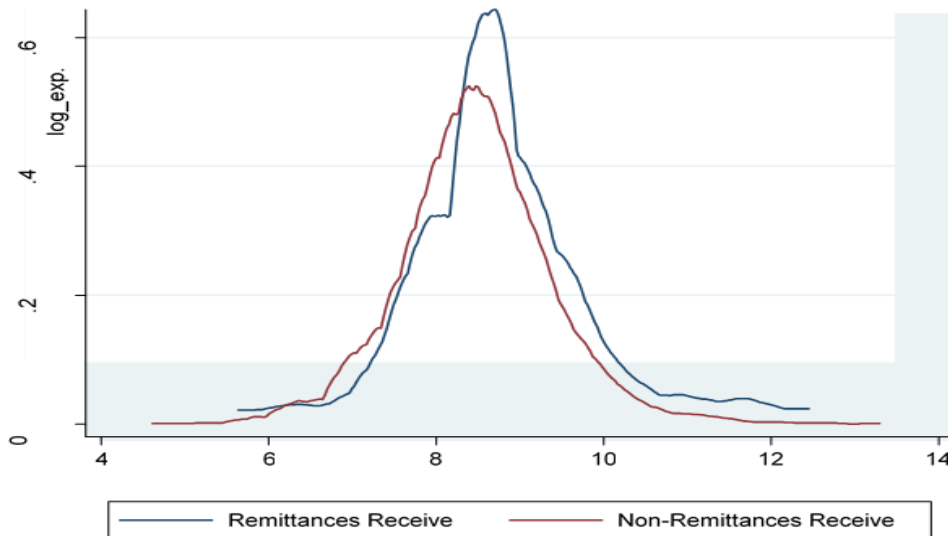


Note: Black dot represent the unmatched pair and cross dot represent matching pair. After using balancing technique of matching the bias is reduced.

APPENDIX C

Overlapping and Balancing Assumptions for Rural Areas

Fig. C1. Propensity of Scores for Remittance Receiver and Non-Receiver in Rural Areas



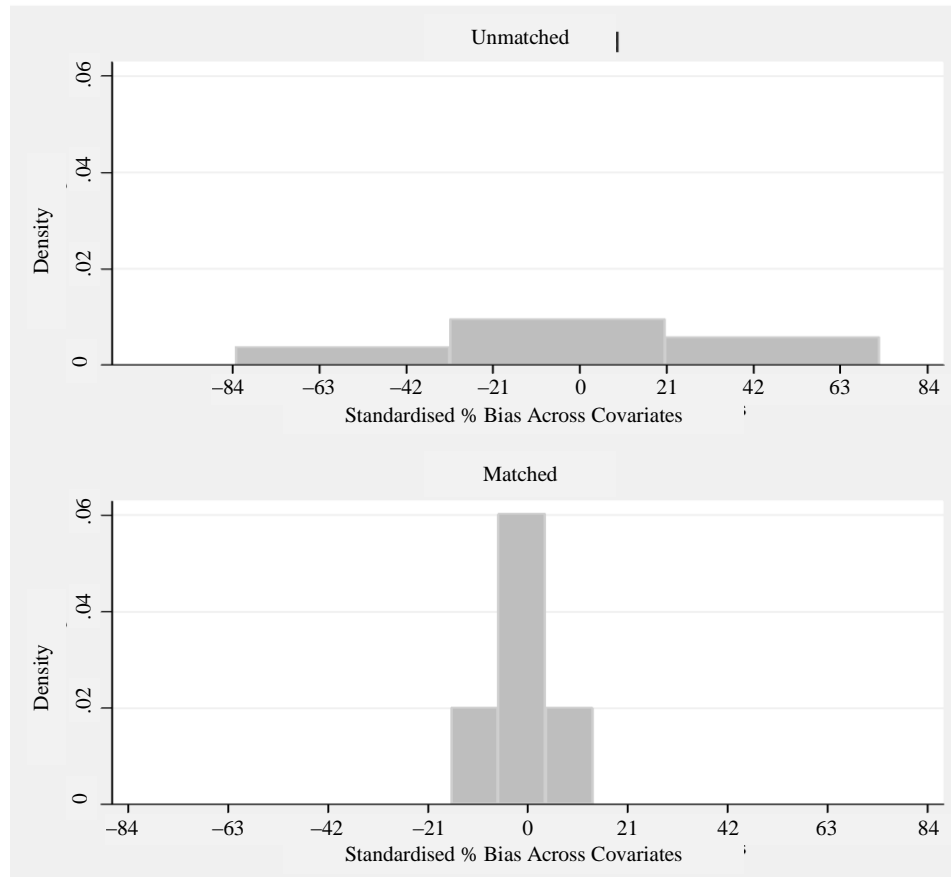
Note: Figure C1 displays that the overlap assumption is satisfied between non-treated and treated.

Table C1
*Covariates between Treated and Non-treated Units of Remittances Receiver
 and Non-receiver (Rural Sample)*

Variable	Mean Treated	Mean Untreated	St. Difference
Head Gender (Male=1)	0.86	0.90	0.088
Household Size	7.22	6.88	0.033
Head age	46.14	45.65	0.033
Education of Head (Illiterate as Reference)			
Grade 1-5	0.12	0.17	-0.125
Grade 6-8	0.10	0.10	-0.009
Grade 9-10	0.11	0.12	-0.029
Grade 11 and above	0.11	0.09	0.099
Province (Balochistan as Reference)			
Punjab	0.45	0.44	0.024
Sindh	0.22	0.25	-0.108
Khyber Pakhtunkhwa	0.27	0.22	0.108

Note: In Table C1, the balance of covariates between treated and non-treated units in rural areas have been shown.

Fig. C2. Standardised Percentage Bias across Covariates in Rural Sample



Note: Figure C2 clearly shows that bias in unmatched is greater as compared to bias in matched.

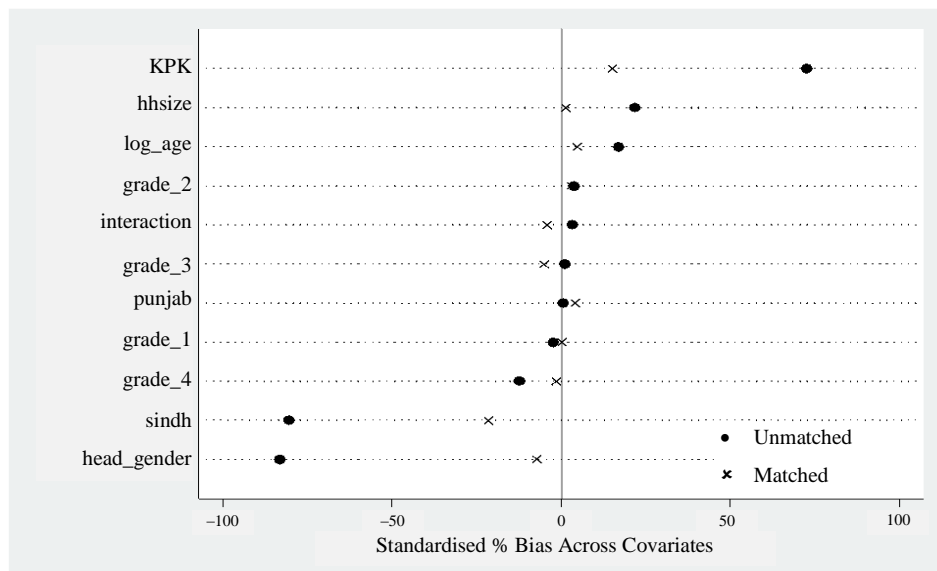
Table C2

Biasness between Unmatched and Matched Variables (Rural Sample)

Variable	Unmatched/Matched U/M	Mean		%Bias	Reduce Bias
		Treated	Control		
Head Gender (Male=1)	U	0.586	0.918	-83.3	91.9
	M	0.586	0.614	-6.8	
Household Size	U	7.686	6.824	21.7	77.7
	M	7.686	7.493	4.9	
Head age	U	48.553	45.457	21.2	82.7
	M	48.553	48.018	3.7	
Education of Head (Illiterate as Reference)					
Grade 1-5	U	0.159	0.168	-2.3	29.2
	M	0.159	0.165	-1.7	
Grade 6-8	U	0.115	0.104	3.7	34.7
	M	0.115	0.108	2.4	
Grade 9-10	U	0.125	0.122	1.1	-94.6
	M	0.125	0.118	2.1	
Grade 11 and above	U	0.058	0.090	-12.4	88.1
	M	0.058	0.061	-1.5	
Province (Balochistan as Reference)					
Punjab	U	0.439	0.436	0.6	33.4
	M	0.439	0.437	0.4	
Sindh	U	0.010	0.268	-80.4	80.0
	M	0.010	0.615	-16.1	
Khyber Pakhtunkhwa	U	0.533	0.203	72.6	81.0
	M	0.533	0.470	13.8	

Note: Table C2 shows the biasness between mean of treated and non-treated households before and after the matching in rural areas. The percentage reduction in bias between treated and no-treated households after matching is shown in the last column.

Fig. C3. Standardise Difference between Matched and Unmatched Variables in Rural Sample



Note: Figure C3, shows the difference between before and after matching for all variables used.

REFERENCES

- Abbas, K., Sabir, H. M., Shehzadi, A., and Abbas, Q. (2014). Impact of workers' remittances on household welfare in district Jhang (A case study of Tehsil 18 Hazari). *Journal of Finance and Economics*, 2(4), 131–135.
- Adams, R. H., and Page, J. (2005). Do international migration and remittances reduce poverty in developing countries? *World Development*, 33(10), 1645–1669.
- Ahmed, J., Mughal, M., & Martinez-Zarzoso, I. (2018b). They earn and send; we spend: consumption patterns of Pakistani migrant households. *International Journal of Social Economics*, 45(7), 1092–1108.
- Amuedo-Dorantes, C., and Pozo, S. (2011). New evidence on the role of remittances on healthcare expenditures by Mexican households. *Review of Economics of the Household*, 9(1), 69–98.
- Arif, G. M. (2004). Effects of overseas migration on household consumption, education, health and labour supply in Pakistan. In Hisaya Oda (ed.) *International labour migration from South Asia*. Chiba, Japan: Institute of Developing Economies and Japan External Trade Organisation.
- Armendariz, B., and Morduch, J. (2010). *The economics of microfinance*. Princeton: The MIT Press.
- Becker, S. O., and Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *The Stata Journal*, 2(4), 358–377.
- Bilquees, F., and Hamid, S. (1981). Impact of international migration on women and children left behind: A case study of a Punjabi village. (PIDE Research Report Series No. 115).
- Bryson, A., Dorsett, R., and Purdon, S. (2002). The use of propensity score matching in the evaluation of active labour market policies. Department for Work and Pensions, United Kingdom. (Working Paper No. 4).
- Caliendo, M., and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31–72.
- Clement, M. (2011). Remittances and household expenditure patterns in Tajikistan: A Propensity score matching analysis. *Asian Development Review*, 28(2), 58–87.
- Dehejia, R. H., and Wahba, S. (2002). Propensity score-matching methods for non-experimental causal studies. *The Review of Economics and Statistics*, 84(1), 151–161.
- Drabo, A., and Ebeke, C. H. (2011). Remittances, public health spending and foreign aid in the access to healthcare services in developing countries. Center for Studies and Research on International Development, Clermont-Ferrand, France. (CERDI Working Paper No. 4).
- Frank, R., Palma-Coca, O., Rauda-Esquivel, J., Olaiz-Fernández, G., Díaz-Olavarrieta, C., and Acevedo-García, D. (2009). The relationship between remittances and health care provision in Mexico. *American Journal of Public Health*, 99(7), 1227–1231.
- Göbel, K. (2013). Remittances, expenditure patterns, and gender: Parametric and semi parametric evidence from Ecuador. *IZA Journal of Migration*, 2(1), 1–9.
- Groot, W. and Maasen Van den Brink, H. (2006). What does education do to our Health? in measuring the effects of education on health and civic engagement: Proceedings of the Copenhagen Symposium, OECD: Paris.

- Ravallion, M. (2003). 'Assessing the poverty impact of an assigned program'. In Bourguignon, F. and L. Pereira da Silva (eds.) *The impact of economic policies on poverty and income distribution*. New York: Oxford University Press.
- Rosenbaum, P. R., and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Ruiz-Arranz, M., and Giuliano, P. (2005). Remittances, financial development, and growth. Washington: International Monetary Fund. (IMF Working Paper No. 05/234).
- Smith, J. A., and Todd, P. E. (2005). Does matching overcome LaLonde's critique of Non-experimental estimators? *Journal of Econometrics*, 125(1), 305–353.
- State Bank of Pakistan (SBP) (2017). *Country-wise workers' remittances*. Karachi: Statistics and DWH Department, State Bank of Pakistan.
- Sudhinaraset, M., Ingram M., Lofthouse, H. K., and Montagu D. (2013). What is the role of informal healthcare providers in developing countries? A systematic review. *PLoS ONE* 8(2):1-12 (e54978).
- Toor, I. A., and Butt, M. S. (2005). Determinants of health care expenditure in Pakistan. *Pakistan Economic and Social Review*, 43(1), 133–150.
- United Nations (2020). *World migration report*. New York: Department of Economic and Social Affairs, United Nations.
- Valero-Gil, J. N. (2009). Remittances and the household's expenditures on health. *Journal of Business Strategies*, 26(1), 119–140.
- World Bank (2019). *Migration and remittances: Recent developments and outlook*. Special Topic: Financing for Development, Migration and Development Brief No. 24. Washington: The World Bank.
- Zweifel, P., Breyer, F., and Kifmann, M. (2009). *Health economics*. Berlin Heidelberg: Springer Science & Business Media.

- Ravallion, M. (2003). 'Assessing the poverty impact of an assigned program'. In Bourguignon, F. and L. Pereira da Silva (eds.) *The impact of economic policies on poverty and income distribution*. New York: Oxford University Press.
- Rosenbaum, P. R., and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Ruiz-Arranz, M., and Giuliano, P. (2005). Remittances, financial development, and growth. Washington: International Monetary Fund. (IMF Working Paper No. 05/234).
- Smith, J. A., and Todd, P. E. (2005). Does matching overcome LaLonde's critique of Non-experimental estimators? *Journal of Econometrics*, 125(1), 305–353.
- State Bank of Pakistan (SBP) (2017). *Country-wise workers' remittances*. Karachi: Statistics and DWH Department, State Bank of Pakistan.
- Sudhinaraset, M., Ingram M., Lofthouse, H. K., and Montagu D. (2013). What is the role of informal healthcare providers in developing countries? A systematic review. *PLoS ONE* 8(2):1-12 (e54978).
- Toor, I. A., and Butt, M. S. (2005). Determinants of health care expenditure in Pakistan. *Pakistan Economic and Social Review*, 43(1), 133–150.
- United Nations (2020). *World migration report*. New York: Department of Economic and Social Affairs, United Nations.
- Valero-Gil, J. N. (2009). Remittances and the household's expenditures on health. *Journal of Business Strategies*, 26(1), 119–140.
- World Bank (2019). *Migration and remittances: Recent developments and outlook*. Special Topic: Financing for Development, Migration and Development Brief No. 24. Washington: The World Bank.
- Zweifel, P., Breyer, F., and Kifmann, M. (2009). *Health economics*. Berlin Heidelberg: Springer Science & Business Media.

Commentary

Machine Learning for Economists: An Introduction

SONAN MEMON

1. INTRODUCTION

Machine Learning (henceforth ML) refers to the set of algorithms and computational methods which enable computers to learn patterns from training data without being explicitly programmed to do so.¹ ML uses *training data* to learn patterns by estimating a mathematical model and making predictions in *out of sample* based on new or unseen input data. ML has the tremendous capacity to discover complex, flexible and crucially *generalisable* structure in training data. Conceptually speaking, ML can be thought of as a set of complex function approximation techniques which help us learn the unknown and potentially highly nonlinear mapping between the data and prediction outcomes, outperforming traditional techniques.²

In this exposition, my aim is to provide a basic and non-technical overview of machine learning and its applications for economists including development economists. For more technical and complete treatments, you may consult Alpaydin (2020) and James, et al. (2013). You may also wish to refer to my four lecture series on machine learning on YouTube [https:// www.youtube.com/watch?v=E9dLEAZW3L4](https://www.youtube.com/watch?v=E9dLEAZW3L4) and my GitHub page for detailed and more technical lecture slides <https://github.com/sonanmemon/Introduction-to-ML-For-Economists>.

ML applications have littered the academic literature and triumphed in industry applications. A case in point is Deep Face, a deep neural network created by Facebook for facial recognition. Another poster child for ML's success is Deep Mind's AlphaGo programme based on neural networks which defeated the world Go champion in 2016. In addition, numerous applications abound in diverse areas such as fraud detection, spam filtering, speech recognition, recommendation systems, medical diagnosis, gene prediction based on DNA sequences in genomics, sales prediction for supermarkets, customer segmentation research, stock market prediction and house price prediction.

During the past few years, economists have also harnessed the power of machine learning in their research. A few applications from recent economic literature include training

Sonan Memon is affiliated with the Pakistan Institute of Development Economics, Islamabad.

¹ML is not identical to Artificial Intelligence (AI). It is more accurate to think of ML as a subset of AI.

²Using Chebyshev polynomials or manual human effort to approximate functions are examples of traditional methods.

neural nets on satellite data to predict local economic outcomes in African countries (Jean, et al. 2016). Cellphone usage data and ML has been used to measure wealth and quantify poverty in Rwanda (Blumenstock, et al. 2015); Bangladesh (Steele, et al. 2017) and to identify ultra-poor households for targeting development aid better in Afghanistan (Aiken, et al. 2020). Larsen, et al. (2021) used text data on news and ML to estimate the impact of news on household inflation expectations.

2. KEY CONCEPTS IN MACHINE LEARNING

Broadly speaking, ML falls under two categories: supervised learning and unsupervised learning. Supervised learning involves training data on inputs X and output Y to learn the true mapping $Y = f(X)$. For instance, estimating the probability of disease given patient characteristics i.e. $P(Y|X)$ requires estimating the conditional probability function. Meanwhile, unsupervised learning does not try to learn $f(X)$ but unearths patterns and associations in the input space X without data on Y .

ML algorithms, when unconstrained are able to estimate an arbitrarily complex function to fit nearly any training data very accurately. However, since our goal is to make out of sample predictions and generalise, we do not want to allow the algorithm to over fit the training data. In order to prevent this *over fitting problem*, we tie the hands of the algorithm through *regularisation*, which constraints its complexity, solving the variance bias trade off. The optimal degree of regularisation is determined by what are known as *tuning parameters* which have to be tuned toward values that optimise out of sample prediction through *cross-validation* for instance. This problem is also referred to as the variance bias trade off and is illustrated in Figure 1. On the left panel, model complexity is too low, which leads to low out of sample variance but very high in-sample bias. Meanwhile, the right panel illustrates the case when there is over fitting with low in-sample bias but very high out of sample variance. Figure 2 illustrates cross-validation, a method in which we partition the data into K folds.

We hold any one fold fixed at a time and train the algorithm on the other folds but use the left out hold for prediction only. We fine tune the *tuning parameters* by minimising the average cross validation error, which helps us solve the variance bias trade off.

Fig. 1.

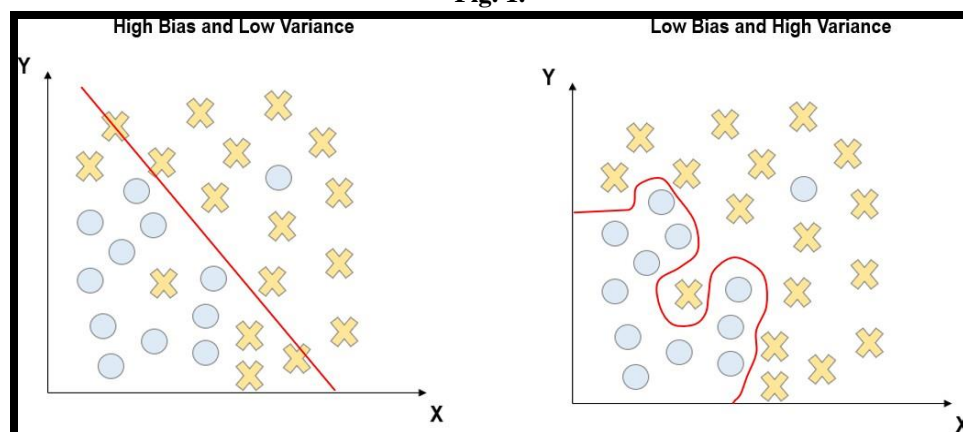
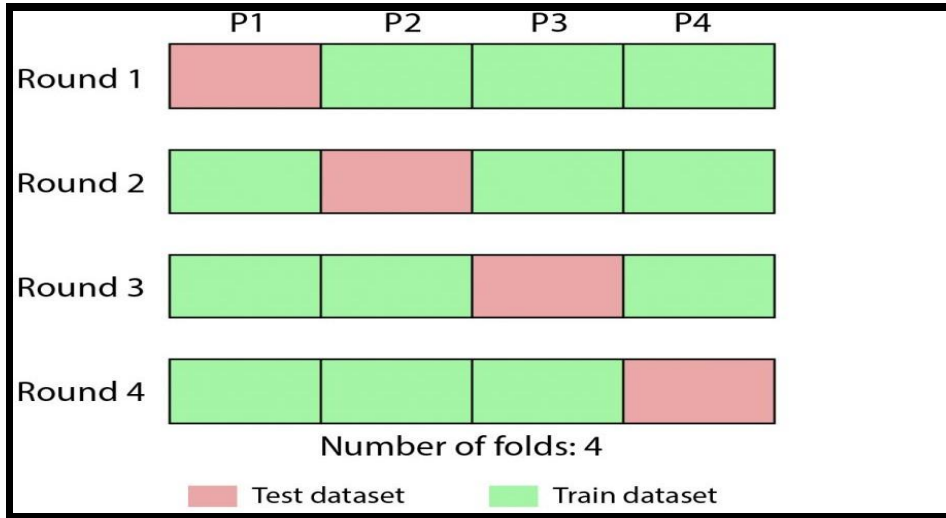


Fig. 2.



For economists, it is essential to keep in mind that the coefficients in ML models do not have causal and policy relevant, structural interpretations unless we impose very strong assumptions on the data generating process. The black box which allows us to learn the true mapping $Y = f(X)$ is still elusive and we do not yet fully understand what is going on behind the scenes. Therefore, the lesson is to look for y^\wedge problems and not β^\wedge problems.

Mullainathan and Spiess (2017), where prediction is the main goal and we are not interested in *identification* of causal parameters in the conventional econometric sense. Having said this, there is now a budding literature which leverages the power of ML to perform causal inference in experimental and observational settings (see Athey, et al. (2015)).

3. OVERVIEW OF METHODS

There is a panoply of different methods available for economists to use, some of which are supervised learning methods for regression problems such as regularised linear regressions (e.g LASSO, elastic nets, ridge regressions), regression trees and random forests, deep learning and neural networks. Meanwhile, algorithms for supervised learning and classification problems include support vector machines and classification trees. If one is focused on unsupervised learning, then K means clustering algorithms and computational linguistics methods such as Latent Dirichlet allocation are some of the options available. In experimental settings, reinforcement learning and multi armed bandits including contextual bandits can help design treatments more optimally. For causal inference using ML, causal forests and other methods can help, especially in identifying heterogeneous treatment effects (see Athey, et al. (2015)). However, in the interest of brevity, I will provide a concise overview of only three methods: LASSO regression, multi-armed bandit problems and computational linguistics. For further understanding regarding ML in economics see Athey (2019) and Athey and Imbens (2019).

4. LASSO

When we are dealing with big data in the sense of large number of covariates and the goal is to make optimal predictions, it turns out that often a relatively small subset of the covariates is sufficient. In order to identify this optimal subset, we can use regularised linear regression such as Least Absolute Selection and Shrinkage Operator (LASSO).³ Formally speaking, LASSO solves the following problem:

$$Blasso = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 \right\} \text{ s.t. } \sum_{j=1}^p |\beta_j| \leq c.$$

A lower level of c , which is a tuning parameter translates into more regularisation or lower complexity.

Making c sufficiently small will cause some of the coefficients to become exactly zero. The remaining set of retained non-zero coefficients are also shrunk toward zero by LASSO.

LASSO is now being frequently used in macroeconomic forecasting as well as in big scanner data for supermarkets and data in neuroeconomics. For example, Kock, et al. (2012) use a Stock and Watson type data set which has 131 macroeconomic time series for macroeconomic forecasting. LASSO throws away many of these variables in prediction problems and retains a small subset which it also shrinks toward zero, favouring sparsity of model specification.

5. MULTI-ARMED BANDITS

The second method that I will discuss is the multi-armed bandit problem, including contextual bandits. Traditionally, experiments were designed by assigning a predetermined number of units to each of several treatments. After outcomes are measured the average effect of the treatment would be estimated using the difference in average outcomes by treatment. This is inefficient since we waste units by assigning them to treatments that are known, albeit with a high degree of uncertainty to be inferior to some of the other treatments. Modern methods for experimentation focus on balancing *exploration* of new treatments⁴ with *exploitation* of returns from treatments that are currently known to work well though these may not be the best ones.

In multi-armed bandits, treatment assignment is *adaptive* and *Bayesian*, updated over time as one keeps on assigning a sequence of incoming units to various treatments. Over time, we essentially estimate the probability of each treatment being the optimal one. We re-evaluate the assignment probabilities after a batch of new observations has arrived in a Bayesian fashion. Figure 3 below depicts an octopus, which has six arms, which correspond to six treatments. Corresponding to each treatment there is a Beta distribution⁵ for payoffs, which is updated over time as the octopus learns about the payoff distribution of all the treatments. As our understanding of payoff distributions improves, we allocate upcoming units to treatments with higher expected returns more often.

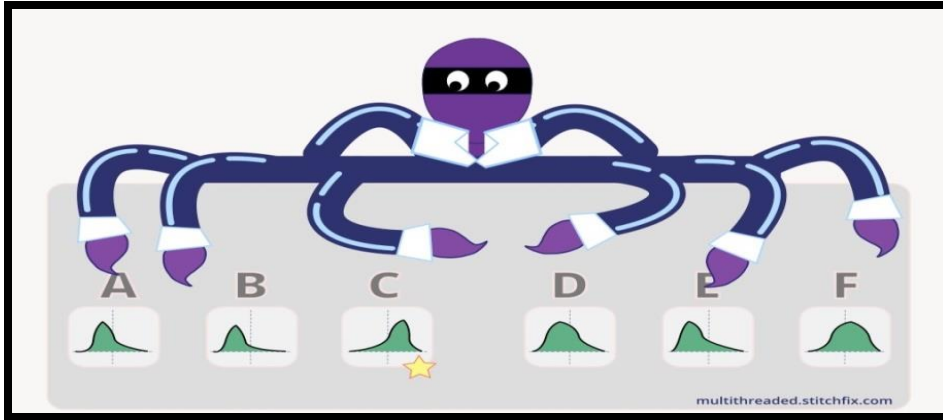
³There are also other methods such as elastic nets and ridge regressions.

⁴This can only be done if we assign units to diverse range of treatments and explore their returns. Obviously, this involves a risk as many of the treatments may have low payoffs.

⁵The Beta distribution is used because of its flexibility and it naturally arises in the binomial case when each arm can return either success or failure.

If all the successive units that arrive are treated as identical, then we have the standard multi-armed bandit problem. However, in experimental settings with humans, there is significant heterogeneity in units, which matters since outcome probabilities vary by unit characteristics. For instance, age, sex and genetic profile is relevant for outcome of drug trials and the probability of finding a job in response to the same labour market intervention will vary across people. When multi-armed bandit problems account for these contextual effects of treatments, they are called *contextual bandits*.

Fig. 3.



For instance, Caria, et al. (2021) use a version of Thompson Sampling algorithm⁶ and contextual bandits for adaptive, targeted treatment assignment in a field experiment for improving job finding rate for Syrian refugees in Jordan. The algorithm balances the goal of maximising participant welfare and precision of treatment effect estimates. Caria, et al. (2021) found that after four months, cash provision has a sizable effect on employment and earnings of Syrians, while some of the other treatments such as information provision and psychological nudge were less effective.

6. APPLICATION OF CONTEXTUAL BANDITS IN EHSAA PROGRAMME

I have currently started work on a project which aims to apply contextual, multi-armed bandit problem to improve design of treatments in the *Ehsaas* programme, Pakistan. This programme includes the BISP⁷ initiative and many other health, economic and education interventions some of which include *Ehsaas Kifalat*, *Nashonuma*, *Tahafuz*, *Ehsaas* undergraduate scholarship programmes, emergency cash transfer programme and many others (see <https://www.pass.gov.pk/home> for details).

I propose that rather than having a priori criteria for assigning a particular *treatment* or mix of treatments/interventions to people with certain demographic and socio-economic characteristics, one could use machine learning to learn how to optimally assign these

⁶Thompson Sampling is a popular algorithm to computationally solve a multi-armed bandit problem. Upper Confidence Bound (UCB) algorithm is another option.

⁷Benazir Income Support Programme.

treatments to maximise human welfare. In order to adaptively update the probability of assigning the various treatments to beneficiaries of *Ehsaas*, one can use algorithms which learn the mapping between individual characteristics and outcomes over time. For instance, consider that there are four possible treatments that are assigned with *a priori* probability of 25 percent each to a certain group of people. A contextual multi-armed bandit can adaptively learn over time which mix of treatments works best for which group of people. Once we have learned from this algorithm, we may be assigning treatment A with probability 60 percent and treatment B with probability 30 percent and the remaining two with only 5 percent probability each to a demographic group with certain features. These optimal treatment assignment probabilities which will vary by individual characteristics cannot be learned without a data driven and machine learning approach.

7. COMPUTATIONAL LINGUISTICS

One big contribution of ML to econometrics is that it makes new forms of data amenable to quantitative analysis: text, images, satellite data, cellphone use data etc. This brings me to my third class of methods, which include algorithms for analysing text data. One method within this class that is extremely influential and has inspired some of the best, recent work in economics on text data is topical modeling of text corpora using Latent Dirichlet Allocation. For a comprehensive survey of literature using text as data in economics see Gentzkow, et al. (2019).

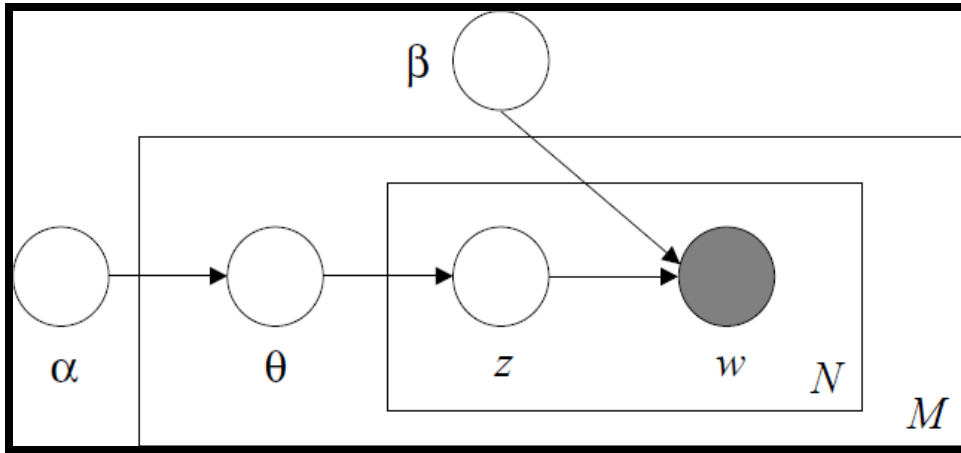
Latent Dirichlet allocation is a three-level hierarchical Bayesian model, also known as a generative probabilistic model (for technical details see Blei, et al. (2003)) for modeling collection of discrete data such as text corpora. In this literature, a document is simply a string of words and a corpus is collection of documents. A topic (z) is a probability distribution over the underlying topics. A word (w) is a probability distribution over topics. For instance, if the topic is about positive sentiment, then words which correlate with positive affect will have higher probability of being associated with this topic as opposed to other topics. LDA allows for topic probabilities to vary across documents, so that we can allow for the fact that some documents such as news articles are more optimistic than others, for instance. We can also allow for multiple topics to co-exist within the same document which allows a richer representation of the diverse information within a document.⁸

LDA first draws a parameter θ for each document in the corpus from the Dirichlet distribution with hyper parameter α , which always returns values from a $K-1$ dimensional simplex when there are K topics. Subsequently, it draws topics for each word in the document from a multinomial distribution with parameter θ . Then, it draws words from the distribution, conditional on topics. Finally, the probability of a document, which is ultimately a distribution over words can also be determined. This hierarchical Bayesian process is illustrated in the following “plate diagram” in Figure 4. The parameters in LDA are estimated using Bayesian methods.⁹

⁸This is unlike previous methods for text analysis such as unigrams and mixture of unigram models Blei, et al. (2003).

⁹Markov Chain Monte Carlo (MCMC) methods, especially Gibbs Sampling methods are used to estimate these models. The original Blei, et al. (2003) paper also proposed an expectation maximisation algorithm which is based on variational inference.

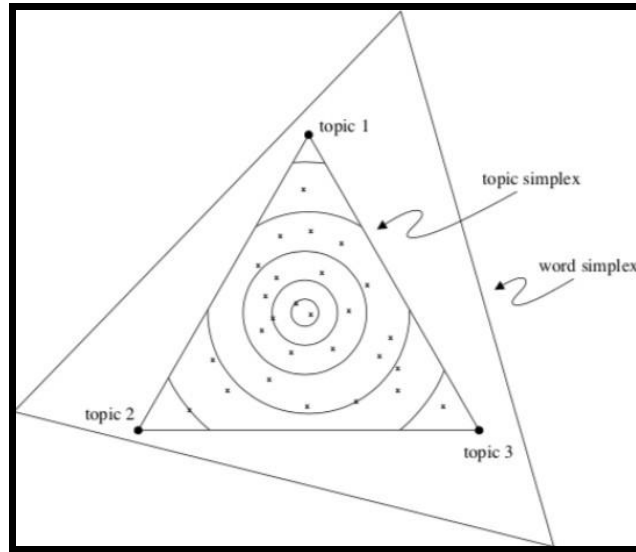
Fig. 4.



Source: Blei, et al. (2003).

Each item or word of a document is modeled as a probability distribution over an underlying set of *topics*. Each topic is in turn a distribution over the underlying set of topic probabilities. This gives rise to a word simplex,¹⁰ where each word is a probability distribution over topics and a topic simplex, which is embedded within the word simplex as shown in Figure 5 below.

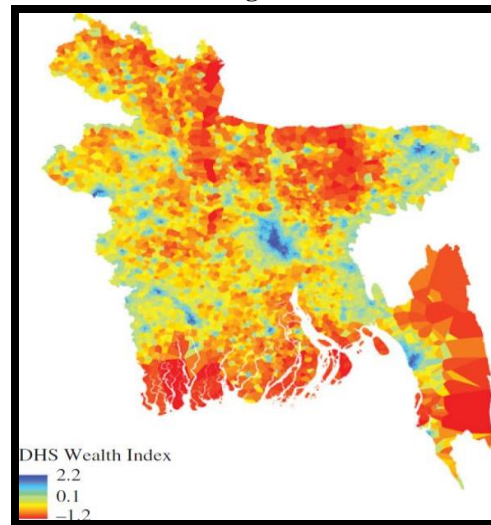
Fig. 5.



Source: Blei, et al. (2003).

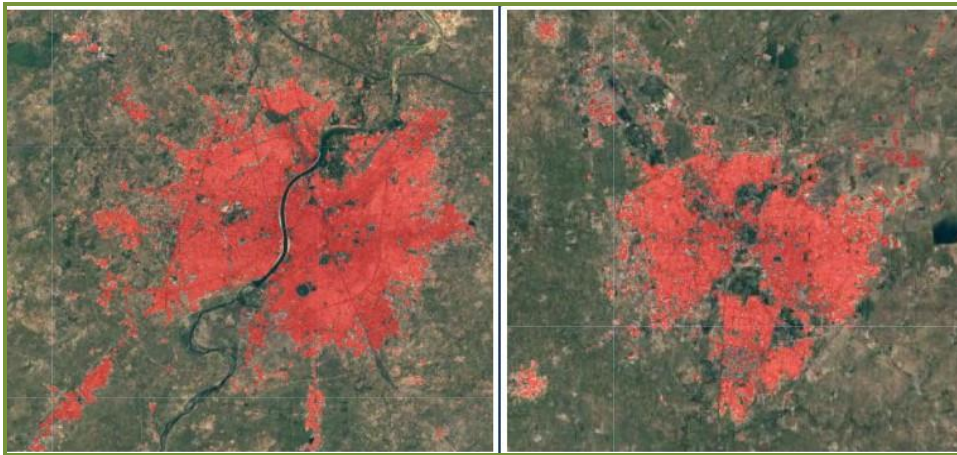
In the context of text modelling, the topic probabilities provide a compact representation of a document. For instance, if we choose to estimate 5 topics based on a collection of documents, LDA will identify these topics based on associations of words in

¹⁰ Note that a simplex is defined by the set of vectors for which it is true that the components sum to one.

Fig. 7.

Source: Steele, et al. (2017).

Goldblatt, et al. (2018) used night-time lights data to “train” for better classification of urban areas in daytime satellite images (Landsat) in the form of built versus non-built areas. They used night-time luminosity data as inputs to predict the probability that a given spatial unit is a built area such as for residential or industrial or commercial purpose. Figure 8 shows a map, which identifies built urban areas in red for a region in India at a highly granular level. Non-built areas may include water bodies and vegetation. This analysis can produce more accurate data on the pace and extent of urbanisation, improving infrastructure development, industrial policy, environmental planning, and land management.

Fig. 8.

Source: Goldblatt, et al. (2018).

Such methodologies enable us to conduct novel and highly granular analysis, which can be updated at low cost and high frequency, addressing the challenges inherent in data scarce environments of developing countries. While it is true that ML systems for

the most part and on their own cannot help us make causal inferences, but with big data, they can enhance predictions, which can automate policy decisions, identify vulnerable populations and regions as well as provide valuable inputs to causal analyses.

9. CONCLUSION

ML has created plethora of new opportunities for economic researchers. It is about time that we should begin to deploy big data and machine learning tools more commonly in academic research and public policy design in Pakistan. These algorithms can improve prediction, enhance the scope of causal inferences, partially compensate for deficiency of rich data by making new data sources exploitable and improve design and efficiency of various public sector programmes in Pakistan. ML can help improve tax compliance, targeting of social protection, make education and health interventions in the *Ehsaas* programme more effective, facilitate in building rich urban profiles and create rich development and security indicators across space more comprehensively among many other payoffs.

REFERENCES

- Aiken, Emily L, Bedoya, Guadalupe, Coville, Aidan, & Blumenstock, & Joshua E. (2020). Targeting development aid with machine learning and mobile phone data: Evidence from an anti-poverty intervention in Afghanistan. In “Proceedings of the 3rd ACM SIG- CAS Conference on Computing and Sustainable Societies” 2020, pp. 310–311.
- Alpaydin, Ethem (2020). *Introduction to machine learning*. MIT Press.
- Athey, Susan (2019). The impact of machine learning on economics. In *The economics of artificial intelligence*. University of Chicago Press, pp. 507–552.
- Athey & Imbens, Guido W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 685–725.
- Athey, et al. (2015). Machine learning for estimating heterogeneous causal effects. (Technical Report).
- Blei, David M., Ng, Andrew Y., & Jordan, Michael I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3, 993–1022.
- Blumenstock, Joshua, Cadamuro, Gabriel, & On, Robert (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, 350(6264), 1073–1076.
- Caria, Stefano, Kasy, Maximilian, Quinn, Simon, Shami, Soha, Teytelboym, Alex, et al. (2021). An adaptive targeted field experiment: Job search assistance for refugees in Jordan.
- Gentzkow, Matthew, Kelly, Bryan, & Taddy, Matt (2019). Text as data. *Journal of Economic Literature*, 57 (3), 535–74.
- Goldblatt, Ran, Stuhlmacher, Michelle F. & Tellman, Beth, Clinton, Nicholas, Hanson, Gordon, Georgescu, Matei, Wang, Chuyuan, Serrano-Candela, Fidel. Khandelwal, Amit K., Cheng, Wan-Hwa, et al. (2018). Using Landsat and night-time lights for supervised pixel-based image classification of urban land cover. *Remote Sensing of Environment*, 205, 253–275.
- Hansen, Stephen, & McMahon, Michael (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, 99, S114–S133.

- James, Gareth, Witten, Daniela, Hastie, Trevor, & Tibshirani, Robert (2013). *An introduction to statistical learning*, Vol. 112, Springer.
- Jean, Neal, Burke, Marshall, Xie, Michael, Davis, W. Matthew, Lobell, David B., & Ermon, Stefano (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301), 790–794.
- Kock, Anders Bredahl & Callot, Laurent, A. F. et al. (2012). *Oracle efficient estimation and forecasting with the adaptive lasso and the adaptive group lasso in vector autoregressions*. School of Economics and Management.
- Larsen, Vegard H., Thorsrud, Leif Anders, & Zhulanova, Julia (2021). News-driven inflation expectations and information rigidities. *Journal of Monetary Economics*, 117, 507–520.
- Mahajan, Aprajit, & Mittal, Shekhar (2017). Enforcement in value added tax: Is third party verification effective? (International Growth Centre Working Paper S-89412-INC-1).
- Mullainathan, Sendhil & Spiess, Jann (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87–106.
- Steele, Jessica E., Sundsøy, Pål Roe, Pezzulo, Carla, Alegana, Victor A., Bird, Tomas J., Blumenstock, Joshua, Bjelland, Johannes, Engø-Monsen, Kenth, De Montjoye, Yves-Alexandre, & Iqbal, Asif M. et al. (2017). Mapping poverty using mobile phone and satellite data. *Journal of The Royal Society Interface*, 14(127).

Why do We Have Less Investment than China and India?

HAFSA HINA

1. INTRODUCTION

Pakistan has experienced macroeconomic instability since the early seventies. Because of the country's persistent macroeconomic uncertainty, savings and private investment have been discouraged, resulting in low aggregate investment and volatile output levels. It has one of the lowest investment-to-GDP ratios that is 15 percent, about half of the South Asian average of 30 percent.

Here we will review the evidence from Pakistan to inform policymaking and local research about

- (1) The investment trends and patterns in the economy and comparison with its regional counterparts.
- (2) The factors which can stimulate the investment.

2. INVESTMENT TRENDS AND PATTERNS

The investment trends and patterns in the economy provide information to understand the abrupt shifts in economic policies implemented by various governments and their effect on the economy. This overview of the economy highlights the problems due to which investment in Pakistan is still uncertain.

Figure 1, shows the trends and patterns of investment (private, public, and FDI), saving, and gross domestic product. Foreign investment is missing that is around 0.6 percent of GDP, It means that foreign investors are not taking any interest in Pakistan. Pakistan trapped in a low-saving, low-investment trap, which limiting its economic potential and long run growth, which is volatile and declining. Private and public invest follow a mixed trends that's why, Figure 1, is further divided into five sub periods contingent upon the trends of private and public investment. So that we know the internal and external factors that cause low investment rate in Pakistan (see Table 1).

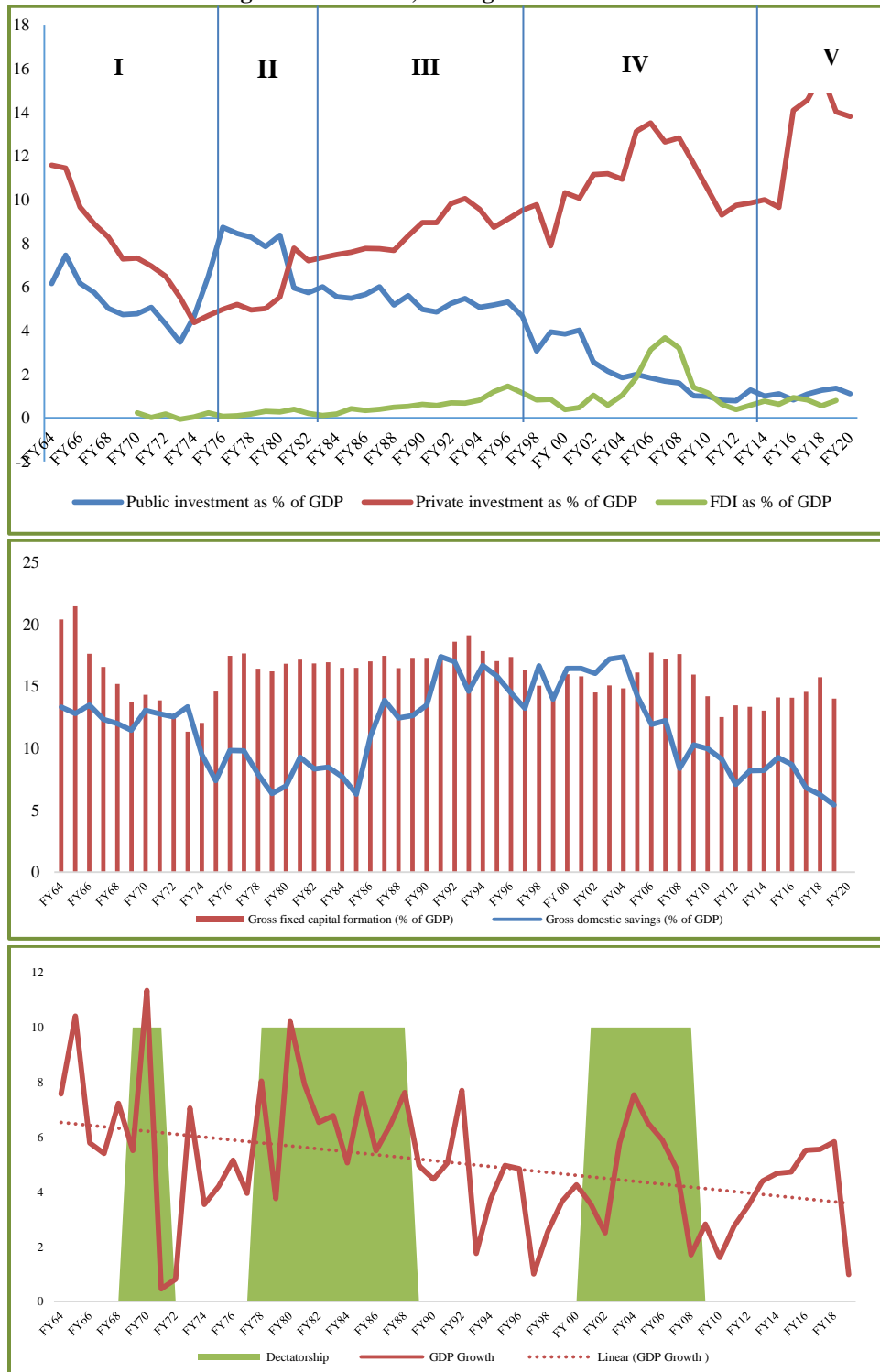
Fig. 1. Investment, Saving and GDP Trends

Table 1

Private and Public Investments Trends in five Sub Periods

Sub Periods	Trends	Average (% of GDP)	
1964-1974	• Private and public investment are declining	Public invest	5.2
		Private invest	8.0
	• Private investment is higher than Public Invest	FDI	0.1
		Saving	12.4
		Total Investment	15.4
		GDP growth	5.9
1975-1980	• Private and public investments are rising	GDP growth	5.9
		Public invest	8.0
	• Public investment dominates the private investment	Private invest	5.1
		FDI	0.2
		Saving	8.0
		Total Investment	6.5
1981-1994	• Private Investment showed raising trend after 1981	Public invest	5.5
		Private invest	8.3
	• Public investment showed a declining trend after 1981	FDI	0.5
		Saving	12.1
	• 1981-1994 investment has fewer fluctuations	Total Investment	17.3
		GDP growth	5.8
1995-2008	• The previous trend continues with huge fluctuations	Public invest	3.1
		Private invest	10.8
		FDI	1.5
		Saving	14.6
		Total Investment	16.0
		GDP growth	4.3
2009-2019	• Seems more volatile period	Public invest	1.0
		Private invest	11.7
		FDI	0.8
		Saving	8.1
		Total Investment	14.1
		GDP growth	3.9

2.1. Internal and External Factors behind the Private and Public Investment of Pakistan¹

1964-1974

- During the 1950s and 1960s, the **private sector was a major channel of industrial investment** in main areas such as banking, insurance, certain basic industries, and trade in major commodities.

¹ This information is taken from Khan and Khilji,(1997), Ahmad (2007), Ahmad and Qayyum (2008, 2009), Farooq (2008), Abbas (2011), Hina (2013), and Lavingia (2016).

- Foreign investment was not allowed in the field of banking, insurance, and commerce (Khan & Kim, 1999).
- The main cause behind the declining trend of public investment was the **1965 Pak-India war and the 1971 Pak-India civil war** in East Pakistan.
- Massive non-development expenditure left an insufficient resource for investment and development purposes.
- In the Ayub regime, private investment was encouraged by concentrating on rapid industrialisation with intensive efforts by the government to promote large scale-manufacturing sector through **expansionary/easy macroeconomic policies** including the facility of tax holidays, tax rebates, and availability of credit.
- The high growth rate during the 1960s was also supported by **foreign aid and assistance** along with policies adopted by the government.
- The disastrous civil war with India in 1971 and large-scale nationalisation adversely influenced the economic growth due to loss of competitiveness in the industry, fall in private investment and shift the trade structure by externalising Pakistan's internal trade.

1975-1980

- Private and public investments are rising and Public investment dominates the private investment.
- The prime reason behind this trend is the **nationalisation policy** of Bhutto's government.
- Large-scale nationalisation of private industrial units and financial institutions.
- Public investment accounted for two-thirds (2/3) of total investment.
- Private investors discouraged due to government nationalisation and the abolition of tax holidays policies.
- In 1977, Zia's military rule came and nationalisation policy was revised gradually and a **mixed economy policy framework** was implemented by the government.
- The 'macroeconomic turmoil' from 1978-80 caused a fall in total investment in many developing economies including Pakistan.

1981-1994

- Private investment showed raising trend after 1981 and public investment showed a declining trend after 1981.
- The government had announced the **industrial policy** in 1984.
- The confidence of private investors grew gradually through the denationalisation of many industrial units and shifting policy towards the pre-1970 policies of investment incentives including five-year tax holidays, import duty reduction on raw material, reduction in the interest rate, and denationalisation of agro-based industries.
- Zia's government adopted several policy measures to attract foreign investment such as **exchange rate liberalisation policy, export processing zone (EPZ)** to encourage export-oriented industries.

- **Foreign investment remained at low levels** during the Zia regime.
- This could be attributed to certain factors including strict licensing and price controls policies, underdeveloped and inefficient financial sector, significant public ownership, high tariffs, and non-competitive trade regime, etc.
- Growth trends were not impressive but improved significantly in comparison to the 1970s period.
- After the Zia regime, the new democratic government came into being and faced the problems of the high budget deficit and worsening balance of payments position and resultantly led the government seeking foreign assistance. Government borrowed from International Monetary Fund (IMF) and started the **Structural Adjustment Program (SAP) of IMF** in the country.
- The establishment of the **Pakistan Board of Investment in 1990** helped generate opportunities for FDI within Pakistan and provide investment services to interested foreign investors. These initiatives placed Pakistan on the International Finance Corporation's list of emerging South Asian stock markets in 1992.
- In 1985 the global economic recession hit the world economy, but it did not significantly impact the Pakistan economy because at that time the economy was not widely opened.
- In 1988 The Gulf crisis originated and many migrants working in the Middle East were sent back home and caused a sharp decline in remittances.

1995-2008

- 1990 decade is important for Pakistan's economy; as **financial sector reforms** were started to promote the private sector and encourage foreign investors.
- The condition of government approval for foreign investment was removed both in industrial and non-industrial sectors.
- Tax holidays, reduction/exemption of customs duty, sales tax, and removal of tariffs, easy visa policy, and certain fiscal incentives were granted to local and foreign investors. Privatisation policy was started in the country.
- Because of these measures, private (domestic) and foreign (direct) investments showed progress and a positive trend was witnessed in both, while public investment followed a slowing and downward trend.
- Total investment and growth in Pakistan showed downturn trend in the wave of Asian financial crisis in 1997 quite similar trends observed in many other Asian countries.
- The **worsened political situation** prevailing in the 1990s and economic sanctions imposed in the wake of **nuclear tests in May 1998**.
- In 1999 the government was collapsed, and Dictator General Pervez Musharraf took over the charge.
- In September 2001, the incident of World Trade Centre (9/11) occurred and it changed the entire scenario of the world and the region as well.
- During Musharraf's rule, the exchange rate remained almost stable at Rs. 60 per US dollar. This stability was contributed to high inflows of remittances and foreign capital inflows to Pakistan.

- The majority of investment policies during the 2000s were based on the notion of privatisation and deregulation. The policy aimed to promote investment in sophisticated, high-tech, and export-oriented industries. While almost all economic operation has been thoroughly open to foreign investment in all other sectors, agriculture, services, infrastructure, etc., with all fiscal and other incentives including loan financing from local banks.
- The level of growth experienced during the early years of the 2000s was impressive, it is important to note that the nation's FDI levels lagged behind the rest of the developing world. In 2007 capital inflows to Pakistan were 4 percent of GDP while average capital inflows to other developing countries were 7.5 percent of GDP this difference stem from an unstable political environment, inadequate infrastructure, and high levels of security risk.
- The global financial crisis of 2008 originated in the US had affected the Pakistan economy like other economies of the world, economic growth declined from 8 percent in 2004 to 2 percent in 2008.
- The global financial crash of 2008 induced further stress on the domestic economy as Pakistani exporters struggled to sell their goods to the nation's largest export market, the United States. Deteriorating diplomatic relations and failure by the Pakistani government to service the nation's debt increased uncertainty over future returns, discouraging foreign firms to invest in Pakistan.

2009-2019

- With the new democratic government in 2008 economy moved from rapid rates of growth to a state of crisis. Real GDP growth slowed sharply and foreign exchange reserves plunged.
- The **shortage of energy and rising security concerns** challenged the nation's capacity to attract foreign investors.
- The PPP government focused on short-term crisis management. Despite reluctance to rely on the IMF, the government turned to the organisation for assistance in November 2008. By accepting IMF financing, the Pakistani government lost an extensive degree of autonomy in designing economic policy and was required to eliminate subsidies in sectors like Energy. Investment growth continued to contract, curtailing public expenditure.
- The investment to GDP ratio stood at 12.5 percent in 2011 at the lowest level. Pakistan's decision to participate in the War on Terror, Pakistan has been perceived as a nation with poor national security. These limitations along with poor government efficacy and political instability have resulted in FDI moving away from Pakistan and towards those developing markets that are less risky to foreign investors.
- Osama Bin Laden's arrest by US Special Forces in 2011 refocused negative attention on Pakistan. Security issues arose as a result of questions about the Pakistani government's ability and willingness to fight terrorism, contributing to a further withdrawal of FDI.

- Following the national elections of 2013, PML-N elected government prioritised to control the domestic energy crisis and curb terrorism for political stability. However, due to constant power outages, poor basic infrastructure, and weak security conditions, Pakistan has been unable to take full advantage of international economic stability and opportunity.
- In 2014, the Board of Investment introduced several additional incentives for foreign investors including 100 percent tax credit for five years on new industries established by June 30th, 2016, as well as credit for investment in infrastructure updates, extensions, and replacements.
- The development of the China-Pakistan Economic Corridor (CPEC), has allowed foreign investment within Pakistan to increase again, especially in the telecommunication, energy, and transportation sectors. Such a partnership is likely to help boost the economy and promote foreign investment within the country.
- PTI government had come into power following the 2018 general elections, it had introduced a strict financial discipline to curtail excessive government expenditure, introduce market-driven exchange rate, and remove large tax exemptions.
- The PTI government followed a liberal foreign investment regime and introduced measures to promote Ease of Doing Business (EoDB) in the country.
- Increased SBP policy rate declined private sector borrowing significantly and an increase in our overall debt.
- Business confidence is low because the government has been uncertain about economic policies and results in high inflation with falling per capita GDP.

We Learnt from Section 2,

Investment remained at low levels due to:

- Unfavourable investment incentive, underdeveloped and inefficient financial sector, significant public ownership, high tariffs, and non-competitive trade regime.
- The market is overregulated by the government
- Government instability and political violence: Pakistan doesn't have a good system to run state affairs. Each political party has a different mindset and policies for Pakistan. Not a single party wants to discuss the issues of Pakistan and what Pakistan needed rather protect their self-interests. Therefore, it can be seen from Table 2, that total investment in democracy is highly volatile/uncertain.

Box 2: Stability Ratio

The stability ratio (standard deviation as a percentage of mean) is used to measure the volatility. Standard deviation is not the best measure of volatility, especially when comparing the different eras when the mean of the series is also different. The stability ratio encounters both mean and standard deviation and provides information about which subsample has a higher standard deviation relative to the mean.

Table 2

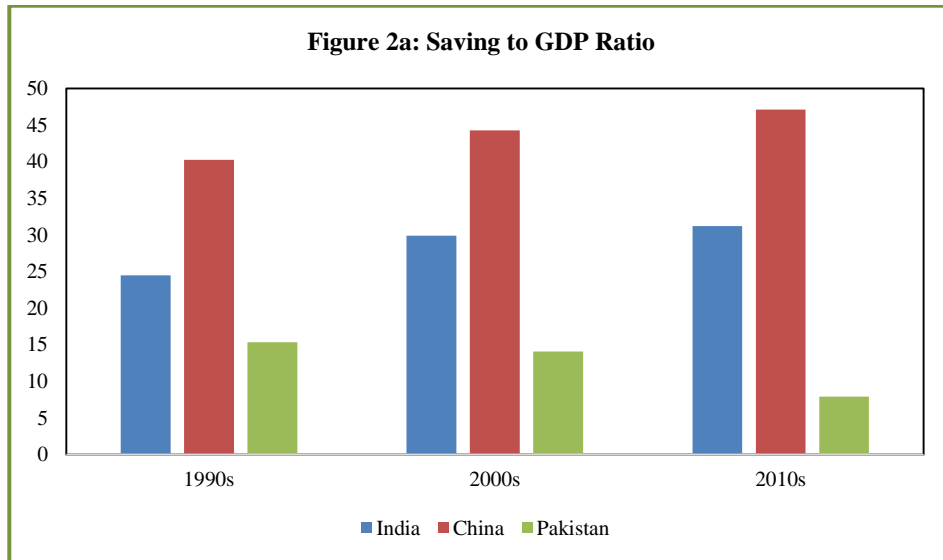
Stability Ratio of Selected Indicators in Military and Democratic Regimes

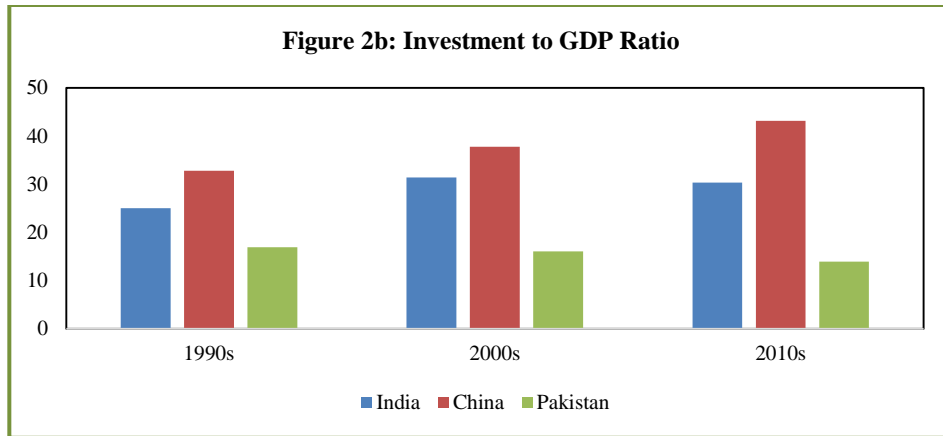
	Pakistan Muslim League (C) 1964-68	Military Dictator- Yahya Khan 1969-71	PPP 1972-77	Military Dictator- Muhammad Zia ul Haq 1978-88	PML(N) PPP 1989-2000	Military Dictator- Pervez Musharraf 2001-08	PML(N) PPP PTI 2009-19
Public Invest % of GDP	0.19	0.04	0.37	0.19	0.16	0.36	0.19
Private Invest % of GDP	0.17	0.03	0.14	0.17	0.08	0.10	0.20
FDI % of GDP	1.09	0.98	1.12	0.40	0.39	0.69	0.37
Saving % of GDP	0.06	0.07	0.21	0.28	0.11	0.22	0.19
Total Investment % of GDP	0.17	0.02	0.19	0.02	0.09	0.08	0.07
GDP Growth	0.31	0.94	0.50	0.25	0.43	0.43	0.43

Source: Author's Calculation.

Low Saving and Investment Rates in Pakistan

In comparison with Indian and China, Pakistan has the lowest level of domestic investment and saving that is 14 percent and 8 percent in the 2010s. Whereas in India investment rate is at 30.3 percent and the saving rate is at 31.2 percent; China has the highest investment rate of 43 percent and saving rate of 47 percent. It seems that Pakistan has zero or no investment and it is trapped in a low-saving, low-investment cycle that limits its ability to grow.



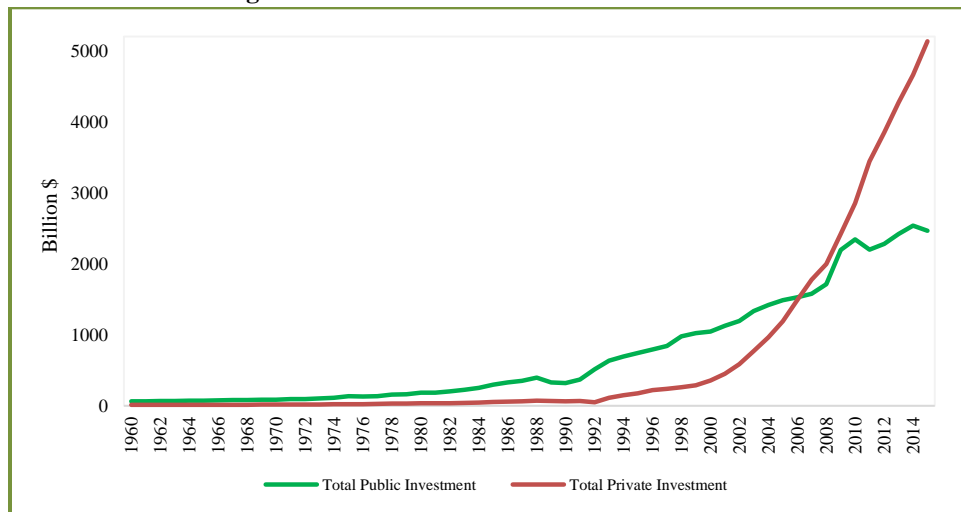


Source: WDI.

Why China is an Investment Champion

- China started its journey as imitators, not as an innovator.
- After the declaration of the ‘Opening-up and Reform’ policy in 1978, China has undergone significant transformations.
- After the 1990s, China adopted new policies that opened up the economy to foreign investment and implemented an unprecedented structure that enabled free enterprise and capitalist ideas to flourish within a socialist framework, resulting in rapid economic and social growth (Ari and Koc, 2020).
- After 1992, private investment per capita in China increased dramatically, eventually surpassing public investment per capita in 2006. From 2006 onward, the Chinese government realised to include innovation as a part of the development strategy.

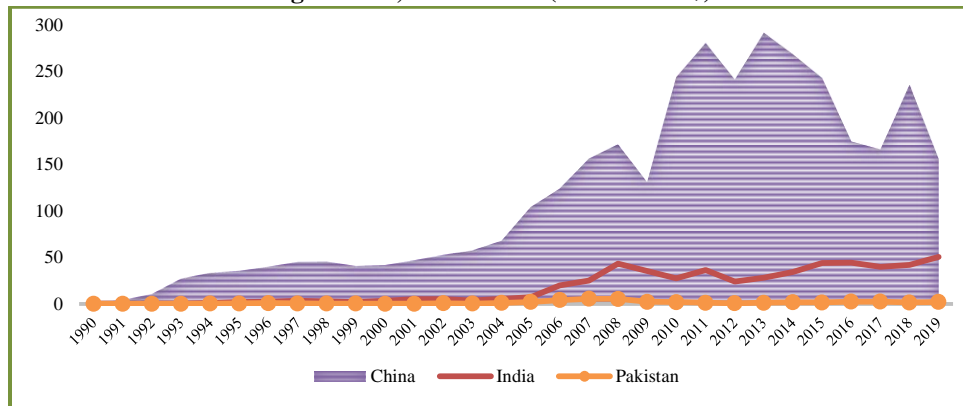
Fig. 3. Public and Private Investment in China



Source: IMF Fiscal Affairs Department.

- China has extraordinary success in attracting FDI over the last 30 years. China's FDI has grown from almost nothing in 1978 to about USD156 billion in 2019. China's great success in attracting FDI under a series of policies since 1978, especially the establishment of a dual capital income tax system (from 1992 to 2008), which provides greater tax concessions to foreign-invested enterprises (Zhang, 2011).

Fig. 4. FDI, Net Inflows (Current US\$)



- China is mostly reliant on domestic investment rather than foreign direct investment. However, FDI contributes not only to the growth of capital, especially in exports, but also to the transfer of excess capital to international markets. This will aid in the transfer of knowledge in the development of human resources (Hina, 2021).
- From the late 1990s, the Chinese government started to strengthen its innovation system and in 2020, innovation in China has taken dramatically on public and business levels. The struggle of three decades on three main factors has led China as a technological leader.
 - The strong role of government.
 - Largest domestic market.
 - Scientific research.

How India Stand out in Investment?

- Local industries established in the late 1960s therefore, the Indian government implemented a more preventive attitude towards FDI.
- In 1973, the new Foreign Exchange Regulation Act (FERA) came into effect, requiring all foreign companies operating in India to register under Indian corporate law with equity capital of up to 40 percent (Kumar, 2003).
- The increase in investments in the mid-1970s was the result of an increase in investments in machinery/equipment. In the early 1970s, more was invested in infrastructure than in equipment. Among the different types of investments, it is the equipment investments that matter most for economic growth (DeLong and Lawrence, 1991).

- The high investment rate is also credited to the financial deepening and expansion of the banking sector in the 1970s and 1980s. Because banks had access to household savings, they could lend to households as well as to businesses.
- In the 1980s, India brought historical changes in its FDI policy. FDI was now considered as a source to earn foreign exchange reserves rather than acting as a supplement to local industries.
- In the 1980s, India made historic changes to its FDI policy. FDI was seen as a means of acquiring foreign exchange reserves rather than as a supplement to local industries.
- The low productivity and inefficiency of local industries are considered to be the result of excessive protection provided to Indian industries from the international market. Such protections made local industries inefficient as compared to other developing countries that pursued liberal FDI policies.
- Foreign direct investment policy in India was reformed by introducing liberal measures. In 1991, India implemented a new economic policy. Since then, India's economy has undergone systematic changes from a highly state-controlled government to a more liberal and outwardly, market-friendly system.
- A series of measures to improve productivity, quality and reduce production costs were gradually introduced (Choudhury, 2018). The lifting of the ban on foreign industries by FERA was a major reform.
- The services sector was opened up to foreign direct investors, especially in the real estate, telecommunications, and banking sectors. In recent years, a series of policy measures have been announced to liberalise the FDI in the country. Gradually, almost all sectors have been opened up to the influx of foreign investment. As a result, India today has one of the most attractive FDI policies in the South Asian region (Sahoo, 2006).
- The informal sector is a big part of the Indian economy. The share of informal employees in the participating labour force is approximately 92 percent. India has taken several steps to address informality, including targeted schemes to promote micro, small, and medium-sized enterprises and legislative measures such as Unorganised Workers Social Security Act, the Contract Labour (Abolition and Regulation) act. And the Workers' Welfare Board. Microfinance has emerged as a means of lending to the informal sector. Since the mid-1980s, the National Bank for Agriculture and Rural Development (NABARD) has been active in a program linking mainstream banks to "self-help groups" (SHGs). Recently, there has been a significant increase in funding for this program in the thirteen priority states that accounts for 70 percent of Indian's poor population. In March 2006, 2.2 million SHGs were connected to regular banks and 33 million poor households had to access to microfinance. NABARD also helps other partner organisations such as NGOs and cooperative banks to promote SHG (Choudhury, 2018).

We Learn from China and India

- Innovation in productivity and management has crucial importance.
- The creation of the linkages between academia and firms for the development of required skills is compulsory which China has created in their innovation process.
- The government should start prioritising technology, science, and innovation as the main pillars for medium and long-term growth as China has done.

- Open the services sector like real estate, telecommunications, and banking sectors to foreign direct investors.
- Overprotection results in low productivity and inefficiency of local industries.
- Equipment investment rather than infrastructure investment matters most for economic growth.

REFERENCES

- Abbas, A. (2011). Essays on investment behaviour in Pakistan, PhD Thesis, PIDE.
- Ahmad, I. (2007). Measuring the effects of public expenditures and macroeconomic uncertainty on private investment: The case of Pakistan, PhD Thesis, PIDE.
- Ahmad, I., & Qayyum, A. (2008). Effect of government spending and macroeconomic uncertainty on private investment in services sector: Evidence from Pakistan. *European Journal of Economics, Finance and Administrative Sciences*, 11, 84–96.
- Ahmad, I., & Qayyum, A. (2009). Role of public expenditures and macroeconomic uncertainty in determining private investment in large scale manufacturing sector in Pakistan. *International Research Journal of Finance and Economics*, 26, 34–40.
- Ari I, Koc M. (2020). Economic growth, public and private investment: A comparative study of China and the United States. *Sustainability*, 12(6), 22–43.
- Choudhury, R. N. (2018). Determinant of foreign direct investment in India's e-commerce sector Singapore: Institute of South Asian Studies, National University of Singapore. (ISAS Working Paper No. 291).
- Choudhury, R. N. (2018). FDI in India and Pakistan: Potential sectors for bilateral investment. *South Asian Survey*, 25(1-2), 129–162.
- DeLong, J. Bradford & Summers, Lawrence H. (1991). Equipment investment and economic growth. *Quarterly Journal of Economics*, 106(2), 445–502.
- Farooq, U. (2008). Economic and non-economic determinants of private investment: A case study of Pakistan, MPhil Thesis, PIDE.
- Hina, H. (2013). Effects of global financial crisis on exchange rate dynamics using various theoretical models: An econometric investigation, PhD Thesis, PIDE.
- Hina, H. (2021). Unpacking investment crisis in Pakistan, Chapter 1, institutions, policy, reform and research: A national narrative-building through webinars at PIDE, PIDE.
- IMF (2017). *FAD investment and capital stock database 2017: Estimating public, private, and PPP capital stocks*. Washington, DC: International Monetary Fund.
- Khan, A., & Khilji, N. (1997). Foreign direct investment in Pakistan: Policies and trends [with comments]. *The Pakistan Development Review*, 36(4), 959–985.
- Kumar, N. (2003). Liberalisation, FDI flows and economic development: An Indian experience in the 1990s New Delhi: Research and Information System for the Non-aligned and Other Developing Countries. (RIS-DP No 65/2003).
- Lavingia, S. (2016). Attracting foreign direct investment in Pakistan, Undergraduate Thesis, Oberlin College.
- Sahoo, P. (2006). Foreign direct investment in South Asia: Policy, trends, impact and determinants Tokyo: Asian Development Bank Institute. (ADB Institute Discussion Paper No 56).
- Zhang, N. (2011). Foreign direct investment in China: Determinants and impacts. The University of Exeter. (Unpublished thesis).

The Pension Bomb and Possible Solutions

MAHMOOD KHALID, NASEEM FARAZ, and BRIG. (RETD.) MUHAMMAD ASHRAF

1. OVERVIEW

Public sector employment remains an attraction for two important reasons: job security and a guaranteed pension (Dixit, 2002).¹ Unlike other countries, Pakistan has not reformed its public sector pension system and has maintained a *pay-as-u-go* defined benefits type pension system which has resulted in build up of unfunded liability for the government. Pakistan practices a legacy pension system where pensioners are paid directly from the revenues as part of the current expenditures. This practice is inherently unsustainable as pension expenditure growing at around 25 percent, cannot be provided from an economy growing at a significantly lower rate. The pension burden is therefore bound to grow, doubling every four-years. In the fiscal year 2018-19, federal superannuation and pension expenditures were almost 78 percent of the value for PSDP expenditures and it increased in FY 2019-20 to 87 percent (463,419 million Rupees and 533,220 million Rupees respectively). The share of pensions as a percentage of current expenditures is also increasing overtime (for FY 2019-20 it stood around 7.6 percent).

Box 1. Pension Driven Fiscal Burden

- Pension expenditures growth: 25 percent.
- Pension expenditures constitutes approximately 80 percent of the value for PSDP expenditures.
- Projected pension values suggest that it will rise to 56 percent of current expenditure in 2050.
- Ageing population, increased medical expenditures, and forced inflation indexing will continue to put pressure on the pension bill.

It is estimated that by 2050, pensions will account for 56 percent of current expenditure. The government will not have the funds for pension expenditure after 8-10

Mahmood Khalid <mahmood.khalid@pide.org.pk> is Senior Research Economist, Pakistan Institute of Development Economics, Islamabad. Naseem Faraz <naseem@pide.org.pk> is Senior Research Economist, Pakistan Institute of Development Economics, Islamabad. Brig. (Retd.) Muhammad Ashraf.

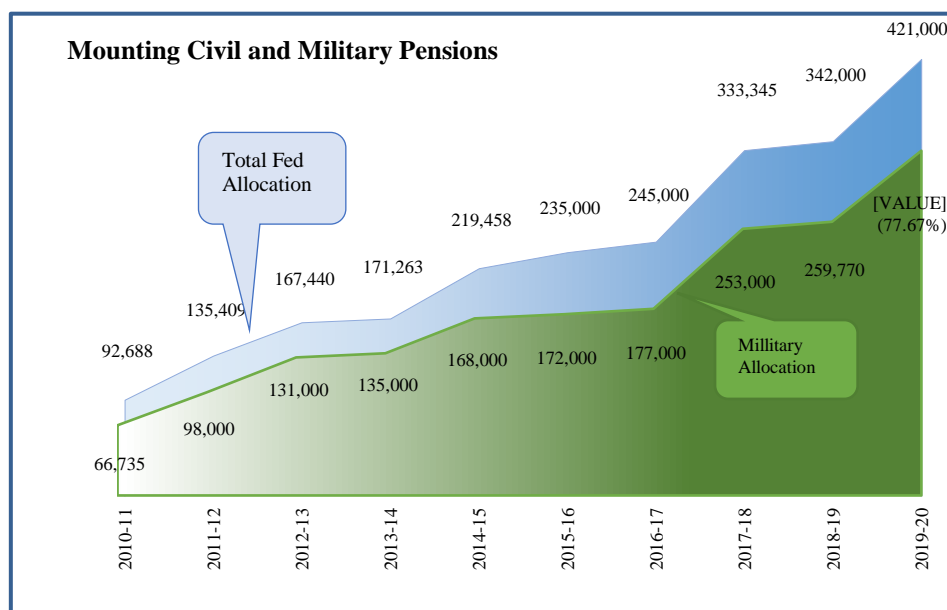
¹Dixit, Avinash (2002) Incentives and organisations in the public sector: An interpretative review. *Journal of Human Resources*, 37(4), Fall, 696–727.

years.² The pension outlay in Punjab Budget equals 95 percent of the revenue of the province. Railway pensions stand at 70 percent of its annual revenue.

In 2003, Pension Review Working Group had recommended the Funded Pension Scheme with Defined Benefit part to be funded by the state and Defined Contributions by the employees. The government, however, did not approve. The fact remains; Pakistan is already in deep crisis with little time left for the transformation of the pensions system.

2. FEDERAL PENSIONS—MILITARY AND CIVILIAN GOVERNMENT EMPLOYEES

It is worth noting that the share of military pensions stands at 75.67 percent of the Federal Pension Budget 2019-20, though the number of the Federal and Military retirees stands at around 1.2mn each. Three times the bigger expenditure of the military is on account of the early retirement of soldiers resulting in pension durations twice the federal employees who retire at 60. However, Civil pension itself is rendering into an unsustainable territory of public financial management. In FY 2019, pension rose to 40 percent of federal public expenditures on salary for civil servants due to longevity, increasing health costs and liberal reforms overtime. Ageing population,³ increased medical expenditures, and forced inflation indexing will all continue to put pressure on the pension bill.



Source: *Economic Survey of Pakistan* (Various Issues).

²Addressing the Pension Liability by Hasaan Khawar, April 18, 2018 <https://tribune.com.pk/story/1681449/addressing-pension-liability>

³According to World Bank, Pakistan had a life expectancy at Birth of 45.3 in 1960, which have improved to 67 years in 2017. Although, it ranked 186th in 249 entries of the data in 2017 which was 166th in 1960. Which reflects that relatively speaking its not that impressive rather it has worsened.

3. NATIONAL COMMISSION FOR GOVERNMENT REFORMS (NCGR) RECOMMENDATIONS: MOVE TO A DEFINED CONTRIBUTION (DC) SYSTEM

The existing system has multiple payout systems to pensioners, is actuarially unfair, biased upward and above all its financing is becoming expensive since it competes for tax payers money. The system is a DB which requires flexibility in terms of contribution, vesting period and other parameters in case of any reform or shock such as increase in life expectancy, cost elements, inflation indexing etc. But since here its a Pay-As-You Go System these options are not considered. This system is unsustainable and world over considered as a least preferred one.

This is not the first incidence, earlier NCGR in its report (Vol-I, 2008) recommended to reform pensions and the recommendation was presented as a guiding principle. It stated that: “The Pension System should be revised from defined benefits, to defined contribution and should be funded.”

NCGR further recommended to only index the pensions for maintaining the purchasing power of pensioners, which lately is being overprovided. The commission proposed a complete road map which required “Complete Analysis of Pension Options and made recommendations guided by the advice given to the government by the Pay and Pension Committee of the Finance Division, Establishment Division and an Actuarial firm/consultant”. This recommendation could not mature so far.

4. PENSION FUND PRACTICES

In the case of armed services, all other countries facilitate soldiers for their second careers. This, on one hand, increases the contributions towards the pension fund, and on the other hand, delays reimbursement of retirement benefits. It results in significant long-term savings at the national level that are mostly invested in the assets which fuel economic growth.

Box 2:Pension Practices in Some Other Countries				
Countries	OECD	EAC	LAC	MENA
	24	10	9	10
Public Sector				
None	3			
DB	15	8	2	10
DC	3	2	—	—
Private Sector				
DB	3	—	—	—
DC	5	5	8	—

Source: World Bank *Pensions Panorama*.

For Civil services, having a DB system which is unfunded creates barriers for entry and exit into the civil services. Turnovers are not possible because of the unrealised pension benefits. If this system is converted into a DC system, then employees would have more leverage for switching across governments/autonomous bodies/ recruitment types and private sector as well. This would help attract qualified and talented individuals in service in various stages of service.

Box 3.⁴ Pension Schemes in South Asian Region		
South Asia	Civil Servants	Special Schemes
Afghanistan	Separate Scheme	Military, Police, Banking sector
Bangladesh	Separate Scheme	—
Bhutan	Main Scheme	Military
India	Separate Scheme	Coal miners, Railway employees
Maldives	Integrated Scheme	—
Nepal	Separate Scheme	—
Pakistan	Separate Scheme	Armed Forces, Police, Railways Employees and others
Sri Lanka	Separate Scheme	Farmers and Fishermen.

However, an actuarial analysis has to be done for identifying the possible benefits which these funds can offer. World over, these pension funds are operated by trusts, private funds managers and even governments themselves. But these funds play a pivotal role in providing for the financing requirements of the economy as well. So, this system would also unleash the growth along with making this sustainable for the government. The role of Pension Funds in the economic growth can be assessed from the fact that around 76 percent of assets in the US Stock Exchange are owned by the Pension Funds.⁵

Pension Funds need to be managed under a strict regulatory regime by a highly competent team of fund operatives and an investment board. Pakistani record in pension funds like EOBI and Volunteer Pension Scheme is well below par.

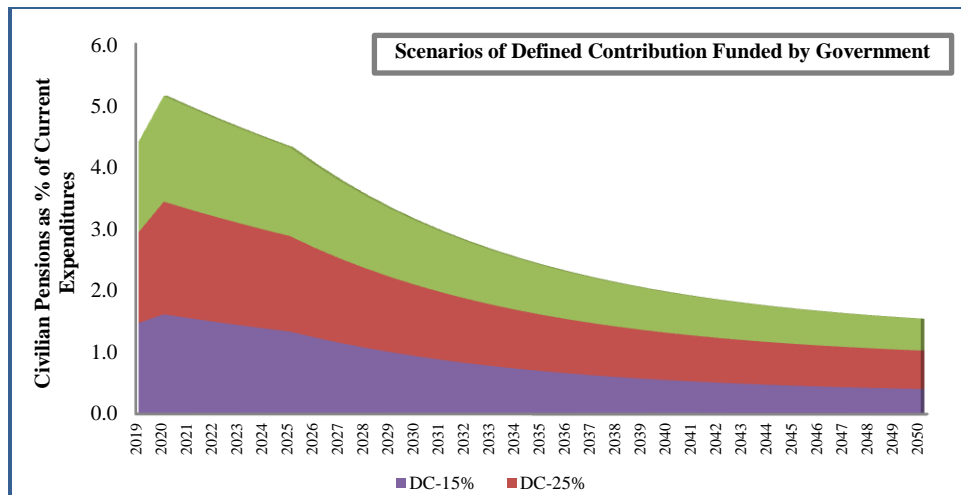
5. CASE STUDY OF FEDERAL CIVIL PENSIONS

Below, we have done projections for a DC system on financing side based on the following assumptions:

- Existing Pensioners would get the same pensions as they are receiving now. Overall volume is freezed, exits from the system would allow inflation indexing for remaining. Using the budget estimates of Civil Federal Pensions for 2019-20, these are projected to remain constant for next five years and then start to decline @ 5 percent per annum due to reduction in exiting pensioners.
- For those in government service assuming emoluments reckonable for pension to be current pay bill along with 30 percent of current allowances. Three scenarios of defined contribution funded by government of 25 percent, 20 percent and 15 percent are developed. Final outlay of DC and ongoing existing pension expenditure as a percentage of current expenditures is plotted.
- Current Expenditures are assumed to increase by 5 percent per annum.

⁴Comparing pension schemes for civil servants across comparable regional countries, we show that there are more than one method being practiced. In some countries, pensions are tailor, made for different professions and in some, it's a universal pension scheme exists irrespective of being a civil servant or otherwise.

⁵*Research Journal of Finance and Accounting* www.iiste.org ISSN 2222-1697 (Paper) ISSN 2222-2847 (Online) Vol.7, No.16, 2016.



This basic exercise identifies that initially total pension outlays increase from 1.5 percent of the total current expenditures to around 1.83 percent, 1.72 percent and 1.61 percent for the three scenarios respectively in the 2020-21 that is the first year of reform. This initial rise tapers down to starting level of 1.5 percent in the 6th, 5th and 3rd year respectively for the three scenarios and then continuously decline and converge around 0.50 percent in thirty years' time period. Thus, making pension payouts to be sustainable in the longrun.

6. THE WAY FORWARD

Pension Funds usually become operative after eight years once sufficient funds have accumulated and investments have started generating revenues. It is, therefore, of critical importance that pensions are transformed into Contributory Funded Pension System to allow time for these funds to become operational before the prevailing system runs out of the stream.

The following policy measures will be required for moving towards the Defined Contribution System of Pensions. Legislation should be enacted for:

- Establishing the Central Pension Development and Regulatory Authority along with an Investment Board.
- Special provisions/exemptions for investments for such projects as housing scheme through the Supplementary Pension Fund.
- Investment in listed and private equities and venture capital.

On the other hand, a Pension Review Working Group (PRWG) should be formed to formulate policies, structural framework, rules, and procedures.

- Composition. The PRWG should comprise the following;
 - Officer with Corporate Exposure.
 - Reps from Existing Armed Services, Federal Services, and Pensioners.
 - Tech Team: Finance and Accounts, Banking, Insurance, Asset Management, Actuarial, Risk Assessment, Corporate Law.

- Consultancy from Pension Fund Managers, Actuaries and Corporate managers if needed.

7. OPTIONS

We have the following broad options:

- Centralised Pension Funds at the Federal and Provincial levels for both; Defined Benefits (Pensions) and Defined Contributions (Retirement Benefits).
- Combination of Centralised Defined Benefits for Pensions but Decentralised Supplementary Fund for the Defined Contributions. An option of the Supplementary Fund for the Military Pension Fund for Defined Contributions, like the Turkish Pension Fund (Oyak), can also be considered.

However, an actuarial analysis has to be done for identifying the possible benefits which these funds can offer.

MoF should urgently initiate Actuarial Study and identify the major issues with government pension schemes, EOBI, and VPS (Voluntary Pension System). It should be aimed at broadening participation of the general public in Pension Funds as seen in western countries.

Books

Nadeem Ul Haque. *Looking Back: How Pakistan Became an Asian Tiger* by 2050. 2017, Karachi: Kitab (Pvt.) Limited.

The book is about development economics and, at the same time, a work of fiction, which predicts the future of Pakistan as a developed country. Though the book is written and conceived on a strong theoretical basis, that is, ‘complexity analysis,’ it remains a fictional work. Because the book is based on complexity analysis, it would fall under the rubric of ‘speculative fiction’, as it attempts to speculate the future development of Pakistan.

As far as speculative fiction is concerned, it is of two types, namely, dystopian and utopian. The book is indeed a ‘utopian speculative fiction’ as it presents Pakistan as a very developed country, contrary to current conditions where Pakistan is striving hard for its survival. Thus, the book equates to the likes of ‘Orwellian’ works, such as “1984”.

Right at the beginning of the book, the author argues that conventional economic analysis has failed to yield any sound analysis of the economy that can be employed for development economics. He further adds that both macroeconomics and microeconomics are inherently insufficient to lead to any meaningful policy analysis and recommendations. Thus, the author presents complexity analysis as an alternative tool for development economics.

According to the author, complexity analysis is inter-disciplinary system analysis without a behavioural model. It does not use the representative, homogenous, and rational agent model. It asserts heterogeneity, diversity, and bounded rationality. It does this in the presence of patchy information, which leads to economic and social systems through repeated actions. Thus, the system does not seek equilibrium but transitions over temporal and spatial milieu.

Simply put, it argues that “the whole can be greater than the sum of the parts.” In short, qualitative change and context are epistemological precursors to an informed analysis, which the complexity analysis provides best to any economic analysis for development economics. Based on the complexity analysis of Pakistan’s political economy, the author broadly presents ten major problems of Pakistan.

- (1) The first problem is the harm caused by foreign aid. Foreign aid usually arrives with political and economic sanctions, which adversely affect the country.
- (2) Rent-seeking by the country’s elite and concentration of power/capital is assumed to be another reason for Pakistan’s woes.
- (3) The greedy institutes that promote bad governance and block all kinds of reforms are also responsible for the sorry state of affairs.

- (4) On nearly all issues, the state policy is fragmented and is not locally owned by the concerned stakeholders.
- (5) The country suffers from a massive brain drain and is responsible for not providing the best minds for the job at hand.
- (6) Constitutional reform is imperative for the country but is not being carried out.
- (7) The Parliament of Pakistan has gone defunct, and it does not function to develop the country.
- (8) The urban centers of the country are mismanaged on a mass scale.
- (9) The colonial mindset of the decision-making public servants also impedes the development process in the country.
- (10) An unnecessary mercantilist approach towards industry also hampers the country's development.

In the next section of the book, a sketch of developed Pakistan in 2050 is presented. In this part of the book, the author proposes strategies for the prosperity of Pakistan. The book deals with all the bottlenecks it has identified in the previous section and presents a remedy for each problem. Finally, the author explains his vision of how Pakistan has developed in 2050.

The first step occurred in 2019 when the state realised its futility and provided seed money to academics and universities to conduct research, free of technical red tapes. This initiated a home-grown quality research initiative that resulted under the umbrella of Research for Pakistan. This was the first step towards rolling a public-driven research agenda in Pakistan, giving birth to 'crowdsourced agenda.'

This initiative resulted in 2022 in forming networks that would debate public issues and guide the public debate on issues, free of donor influence, promoting enlightenment, common knowledge, democratic debate among the citizenry, and, last but not least, the ownership of change among all stakeholders. By 2025, these networks were being heard by the parliament as well. The reforms from this year till 2030 were significant works of this research initiative. This period of Pakistan's history is called enlightenment by the author.

The author argues that this research initiative and the associated networks started to question the notion of 'development first' and advocates the notion of 'build state for development.' Thus, the country's focus shifted, and all efforts were directed towards building the state first so that development could follow. Thus, all state departments, such as police, internal affairs, and others, were restructured. Government performance is now judged by the security of life and property and the quality and speed of justice.

The author next discusses the constitutional amendments carried out to shape and build the state for development. He presents nine constitutional changes backed by the networks in the country: electoral requirements are amended to do away with the over-representation of elite families in the parliament; the arbitrary powers of the prime minister are curtailed, and the cabinet is made more powerful; the budget is now taken seriously due to constitutional amendments; practical rules for an independent judiciary are now in place; a strong local government system is in place; the civil service, the police, public service delivery agencies, and regulatory bodies are independent now and depoliticised; the four provinces have been abandoned, and now the country consists of

20 states of the equal population; and the census is mandatory now, and it is held regularly through constitutional changes and is tied with resource allocation to states.

Moreover, the author argues that the colonial pyramid of authorities is now dismantled, and 'unified pay scales' have been abandoned and replaced with good pay but without privileges, thus making public service more efficient and localised. The author argues that with colonial practices gone, public service made a more responsive and effective decentralised system of governance, the cities are being managed appropriately now, making them engines of growth and, ultimately, developing the country.

The industries have also been reinvented. The author argues that the sugar, cotton, cement, and banking sectors have been overhauled in the presence of strong regulatory authorities, free of political influences. The local markets are flourishing, and the mercantilist approach has been abandoned.

The author also argues that the country is secular now and not an Islamic state. The Mullah (Islamic Cleric) has been regulated, genuine Islamic scholars are being employed to counter Mullah, and the public debate is backing this effort. The reform, led by networks, has successfully placed a regime where even Mullah is properly licensed and does not oppose modernity.

As a result of all the reforms mentioned above, the author argues that the following results have been achieved, which have made Pakistan the Asian Tiger in 2050. Now Pakistani state works on system development and not on sectoral development. Now the state taxes the inheritance and not everything else. Entrepreneurship is thus emerging in Pakistan and helping the country develop. Now opportunities are being provided to the poor due to strong systems the state has put in place. Now the development is public-centric and not industry-centric, as was the case in the past when the poor were ignored and industry, such as the automobile industry, was favoured. Education is now in demand because people know that it is the only way to prosper in Pakistan, and family ties would not give them better prospects. Due to planned urbanisation and inheritance reforms, females are now included in the mainstream and are meaningfully contributing to the country's development.

While listing down the problems of Pakistan, the author has ignored the problems arising out of civil-military relations in the country. A complete absence of this problem from the 'problem tree' constructed by the author merits a revisiting exercise by contemporary researchers for further complexity analysis and to devise recommendations accordingly. Another issue missing in this analysis is the issue of Kashmir. This issue has been draining finances from Pakistan's coffers for a long now. The public agenda set by people, as presented by the author, has completely ignored this vital issue of Pakistan's politics and defense strategy.

Though the author focuses on cities and agricultural reforms, there is a dire need to include discourses on rural development in the book. This is necessary for the complexity analysis to be holistic and inter-disciplinary, as claimed by the author himself. Last but not least, as narrated by the author, in 2050, the UN appoints a 'Pakistan Transformation Commission' to study Pakistan as a successful case study of development for other developing countries. Strangely, the author, who is in favour of multi-disciplinary studies and complexity analysis, has chosen to put only economists as the members of this commission.

Overall, the book is an excellent read and fascinating as utopian fiction, at par with global literature of this kind. It is strongly recommended for all the readers who have an interest in contemporary issues of Pakistan. The book's readability is excellent, and it can be comprehended by all people, especially those with a background in social sciences.

Muhammad Haroon Hussain *and* Abedullah

Pakistan Institute of Development Economics,
Islamabad.



RESEARCH FOR SOCIAL TRANSFORMATION AND ADVANCEMENT

2021 RASTA Competitive Grants Programme for Policy-oriented Research

The Pakistan Institute of Development Economics (PIDE) has launched a multi-year competitive grants programme for policy-oriented research in Pakistan titled ***‘Research for Social Transformation and Advancement’*** (RASTA) under the Public Sector Development Programme (PSDP) of the Ministry of Planning, Development and Special Initiatives, the Government of Pakistan. RASTA’s mission is to develop a research network of academia, think tanks, policymakers, practitioners and other stakeholders across Pakistan producing high-quality, evidence-based policy research to inform Pakistan’s public policy process.

There will be six rounds of the Call for Research Proposals. The first call was in October, 2020, and the second one would be announced in the first quarter of 2021. All updates will be published on PIDE/RASTA website from time to time. In pre-submission engagements webinars and workshops are scheduled to guide potential applicants. For more details and guidelines related to RASTA programme, eligibility, application process and updates, please visit PIDE/RASTA website and follow us on Twitter.

Call for the third round coming soon.

RASTA Project Management Unit

Pakistan Institute of Development Economics,
Adjacent Quaid i Azam University Campus, P.O. Box 1091,
Islamabad 44000, PAKISTAN.

Email: rasta@pide.org.pk | URL: www.pide.org.pk/rasta | Twitter: @RASTA_PIDE

PAKISTAN SOCIETY OF DEVELOPMENT ECONOMISTS

Registered Office:
Pakistan Institute of
Development Economics,
P.O. Box 1091,
ISLAMABAD – 44000

MEMBERSHIP FORM

Dated

FOR OFFICIAL USE

Membership No.

PARTICULARS OF APPLICANT

NAME		DATE OF BIRTH
PRESENT EMPLOYMENT		
DESIGNATION	ORGANISATION	TELE. NO.
ORGANISATION ADDRESS		
HOME ADDRESS		TELE. NO.

ACADEMIC QUALIFICATIONS (Give details of highest degree(s) obtained)

DEGREE	YEAR	UNIVERSITY

EXPERIENCE IN YEARS

TEACHING	RESEARCH	OTHER (Specify)

PUBLICATIONS. Give particulars of articles, monographs & books. (Attach additional Sheet(s) if necessary)

S.NO.	Title	Year

MEMBERSHIP FEE. Rs 20 00.00 (in Pakistan), US\$ 200.00 (overseas members). Please send a crossed cheque to the Pakistan Society of Development Economists.

YEAR(S) FOR WHICH PAID	AMOUNT	DATE	DRAWN ON (Name of Bank & Branch)
CHEQUE NO.			

Please mail to:
Secretary,
Pakistan Society of Development Economists,
PIDE, P.O. Box 1091,
ISLAMABAD – 44000
Pakistan

Signature of Applicant

FOR OFFICIAL USE

DATE APPLICATION RECEIVED	THROUGH (Name of Proposers)	MEMBERSHIP FEE RECEIVED AS ABOVE
	1.	
	2.	

☐

Application Accepted

☐

Application Rejected

REMARKS

PARTICULARS OF MEMBERSHIP

RENEWAL DATE	RENEWED YES/NO	RENEWAL DATE	RENEWED YES/NO	RENEWAL DATE	RENEWED YES/NO

Dated

Secretary

For conditions of eligibility for Membership, see reverse.

**EXTRACTS FROM THE CONSTITUTION OF THE PAKISTAN
SOCIETY OF DEVELOPMENT ECONOMISTS**

ARTICLE 5

- 5.3 *Membership:* There shall be a select category of Members of the Society. The minimum criteria of eligibility for election as Member of the PSDE are:
- (a) Previous *ex-officio* membership of the society;
 - or
 - (b) Master's degree in Economics, Business Administration, Public Administration, Agricultural Economics, Statistics Econometrics or Economic Demography and/or an evidence of proven scholarship in these areas of specialisation.
- 5.4 *Election of Member:* All persons satisfying the minimum eligibility criteria as specified in Article 5.3 may apply for Membership on the prescribed form after having their nominations duly proposed and seconded by any two *Ex-officio* Members/Members of the Society, provided that no such application shall be required of a former *Ex-officio* Member of the Society who may enrol as Member on payment of the prescribed fee at the invitation of the Council. Election to Membership shall be decided by a simple majority of the Executive Council at a constitutionally valid meeting.

THE PAKISTAN DEVELOPMENT REVIEW

Editor: **Nadeem Ul Haque**

The Pakistan Development Review is an internationally refereed journal published regularly by the Pakistan Institute of Development Economics since 1961. The journal focuses on economics and related social sciences and welcomes theoretical and empirical contributions in relevant disciplines with a particular emphasis on Pakistan's socio-economic issues. The journal is published on a tri-annual basis. The journal's editorial and advisory boards consist of more than 18 renowned scholars in the fields of economics and related social sciences. The actively participate in refereeing the papers and also render valuable advice on other related matters.

AIM AND SCOPE

The aim of the journal is to encourage original scholarly contributions that focus on a broad spectrum of development issues using empirical and theoretical approaches to scientific enquiry. With a view to generating scholarly debate on public policy issues, the journal particularly encourages scientific contributions that explore policy relevant issues pertaining to developing economies in general and Pakistan's economy in particular.

ABSTRACTING AND INDEXING

The Pakistan Development Review is indexed and/or abstracted in the EconLit, Scopus, CAB abstracts, Ekistic Index of Periodicals, etc.

THE PAKISTAN DEVELOPMENT REVIEW

SUBSCRIPTION FORM

Kindly enter a subscription to *The Pakistan Development Review* for the Year(s)

in the name of

Address:

.....

.....

The *Review* should be supplied by surface/air mail. A bank draft for the sum of Pak. Rupees/US \$ is enclosed to cover the above subscription.

Signature

Date:

1. For Subscription rates, please see the inside of the back-cover of *The Pakistan Development Review*.
2. Please address your order to: Chief, Publications Division, Pakistan Institute of Development Economics, Post Box 1091, ISLAMABAD – 44000, Pakistan.

E-mail: publications@pide.org.pk

pide@pide.org.pk

Website: <http://www.pide.org.pk>

INSTRUCTIONS FOR AUTHORS

1. All manuscripts submitted for publication should be in English. All submissions, or queries, should be sent by email to: pdr@pide.org.pk. A submission implies that the research work has not been published previously, that it is not under consideration for publication elsewhere and is approved by all authors. The journal also has the policy to verify the originality of the submissions through originality detection service.
2. Each request for a book review in the journal must be accompanied by one copy of the relevant book, which should be submitted to: The Editor, *The Pakistan Development Review*, Post Box 1091, Islamabad-44000, Pakistan.
3. Manuscripts will be accepted for consideration on the understanding that they are original contributions to knowledge in social science fields.
4. All articles should be organised into the following sections: (i) Abstract of 150 words highlighting major contribution and summary of findings followed by *JEL* classification and at least six Keywords, (ii) Introduction covering the hypotheses, objectives of the work, adequate background and literature review highlighting the key gaps in the literature and how the research fills those gaps, (iii) Data and Methodology, (iv) Results and Discussion, and (v) Conclusions and Policy Implications. Sub-sections should carry clear and distinct sub-headings.
5. Each manuscript should be typed single-spaced in times new roman font size 12 (MS WORD) on one side of quarto sheets, and should carry a margin of an inch and a half on the left-hand side of the typed page and of at least an inch on each of the remaining three sides. The total word count of the manuscript should be between 6000-8000 words.
6. The first page of the manuscript should contain: the self-explanatory title of the paper, the name(s) of author(s), and a footnote giving the current affiliation of the author(s), funding agency (if any) and any acknowledgements.
7. As a courtesy to referees, detailed derivations of the main mathematical results reported in the text should be submitted separately along with the articles.
8. Tables for the main text and each of its appendices should be numbered serially and separately. The title of each table as well as the captions of its columns and rows should be clearly expressive of the contents. The source of the table should be given in a footnote immediately below the line at the bottom of the table; but, unlike other footnotes, which must be numbered consecutively, it should not be numbered.
9. Graphs should be sent in editable form and not as pictures. They should be presented in a way that is best suited for black and white printing.
10. Footnotes should be numbered consecutively. Each appendix and each table should have a separate set of footnotes.
11. All references should be arranged on APA style which should be organised through electronic referencing management softwares such as Mendeley and EndNote.
12. The author(s) of each article will receive complimentary copies of *The Pakistan Development Review* in which the relevant contribution appears.
13. Any change in the names of the author(s) after the initial submission is not allowed. Author(s) should make sure to list the names of all contributors, their order and corresponding author before submission.
14. The Journal strictly follows all ethical considerations. At the time of submission, the author(s) are required to disclose potential "conflict of interests" that could inappropriately influence (bias) their work.
15. The journal has no processing/publication fee.

SUBSCRIPTION RATES

PAKISTAN

Annual Subscription

(includes three Regular Issues of the Journal plus the
Papers and Proceedings (Special Edition) of
approximately 700 pages)

	Surface Mail	Air Mail
Individuals	Rs 300.00	Rs 450.00
Institutions	Rs 450.00	Rs 600.00
Per copy of the Regular Issue	Rs 125.00	Rs 150.00
<i>Papers and Proceedings</i> Issue only	Rs 400.00	Rs 500.00

Students

Annual Subscription	Rs 100.00	Rs 250.00
Per copy	Rs 30.00	Rs 65.00

FOREIGN

Annual Subscription

Individuals	US\$ 100.00	€ 85.00	US\$ 125.00	€ 105.00
Institutions	US\$ 150.00	€ 125.00	US\$ 200.00	€ 170.00
Per copy of the Regular Issue	US\$ 40.00	€ 35.00	US\$ 55.00	€ 45.00
<i>Papers and Proceedings</i> Issue only	US\$ 100.00	€ 85.00	US\$ 150.00	€ 125.00

“How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it. Of this kind is pity or compassion, the emotion which we feel for the misery of others, when we either see it, or are made to conceive it in a very lively manner. That we often derive sorrow from the sorrow of others, is a matter of fact too obvious to require any instances to prove it; for this sentiment, like all the other original passions of human nature, is by no means confined to the virtuous and humane, though they perhaps may feel it with the most exquisite sensibility.....

Of all the calamities to which the condition of mortality exposes mankind, the loss of reason appears, to those who have the least spark of humanity, by far the most dreadful, and they behold that last stage of human wretchedness with deeper commiseration than any other. But the poor wretch, who is in it, laughs and sings perhaps, and is altogether insensible of his own misery. The anguish which humanity feels, therefore, at the sight of such an object, cannot be the reflection of any sentiment of the sufferer. The compassion of the spectator must arise altogether from the consideration of what he himself would feel if he was reduced to the same unhappy situation, and, what perhaps is impossible, was at the same time able to regard it with his present reason and judgment.”

The Theory of Moral Sentiments by Adam Smith (1759)

Selected by Durr-e-Nayab

The logo consists of the letters 'PDR' in a bold, serif font, centered within a light blue square.

Pakistan Institute of Development Economics

Quaid-i-Azam University Campus, Islamabad, Pakistan

E-mail: pide@pide.org.pk

Website: <http://www.pide.org.pk>

Fax: +92-51-9248065