

IMF Paper on Policy Analysis and Assessment

© 1998 International Monetary Fund

This is a *Paper on Policy Analysis and Assessment* and the author(s) would welcome any comments on the present text. Citations should refer to a *Paper on Policy Analysis and Assessment of the International Monetary Fund*. The views expressed are those of the author(s) and do not necessarily represent those of the Fund.

PPAA/98/3

INTERNATIONAL MONETARY FUND

IMF Institute

Taxing Agriculture in Pakistan

Prepared by Mahmood H. Khan and Mohsin S. Khan¹

May 1998

Abstract

Pakistan's economic performance has been marred by persistently high fiscal deficits which have resulted largely from the inability to raise sufficient revenues. Agriculture is a dominant sector of the economy, but generates very limited revenue from direct taxes on agricultural producers. The paper first reviews the history of attempts to tax agriculture in Pakistan, the purpose of which is to underscore the difficulty in adopting and implementing a sound agricultural tax policy. It then examines various alternatives with regard to the taxation of agricultural land and incomes. Finally, it presents the outlines of a full-fledged agricultural income tax that could be implemented within the Constitutional framework in the next three to five years.

JEL Classification Numbers: E62, Q18

Keywords: Pakistan, Taxation, Agriculture

Author's E-Mail Address: mkhan@sfu.ca
mkhan@imf.org

¹Mahmood K. Khan is Professor of Economics, Simon Fraser University, British Columbia, Canada. The authors thank Paul Chabrier, Abbas Mirakhor, Vito Tanzi, and Neven Mates for their helpful comments on an earlier draft. Any remaining errors are the responsibility of the authors.

I. INTRODUCTION

One of the striking features of the Pakistan economy has been the inability of successive governments to maintain fiscal discipline on a sustained basis. Since 1971, the year of the breakaway of East Pakistan (now Bangladesh), the government budget has always been in deficit. Whenever this deficit has been large, the growth rate has faltered, inflation has climbed higher, and most importantly, the economy has faced a widening current account deficit that has culminated frequently in a balance of payments crisis.² In the 1990s, the fiscal problem once again became acute, with the budget deficit fluctuating around 6 percent of GDP (Chart 1). As expected, this was accompanied by poor economic performance, characterized by a slowing down of growth to an average of 4 percent per year, as compared to an average annual rate of 6 percent in the 1980s. Inflation also moved to double-digit levels, averaging 12 percent a year. Finally, there was a sharp deterioration in the current account balance, and foreign exchange reserves fell to only two weeks of imports at the end of 1996.

Eliminating the structural fiscal deficit has been the chief preoccupation of all governments in Pakistan, although the efforts to do so have been fairly unsuccessful. In the 1990s, total government revenues stagnated at around 17 percent of GDP, despite a number of revenue-enhancing measures contained in each annual budget (Chart 2). In 1996/97, the revenue ratio dropped to a little over 15 percent, one of the lowest levels in the past

²See Khan (1992).

three decades.³ While the ratio of total expenditure to GDP was cut in the 1990s to offset sluggish revenues, the scope for major reduction is limited. Two items of government expenditure--interest payments and military spending--together comprise about one-half of total expenditure and are generally inviolable to budgetary cutbacks. Thus, increasing revenues has to be the key in any meaningful fiscal adjustment.

The poor revenue picture is the result of a narrow and inelastic tax system that relies heavily on indirect taxes and import taxes (Chart 3). Taxes on income and profits account for only a quarter of total revenues. Furthermore, there are widespread tax exemptions and evasion, which together with poor tax administration ensure that actual revenues consistently run below budget projections. Finally, the income tax system is highly inequitable, with the burden falling mainly on the business sector and salaried workers.⁴ In this context, it is notable that agriculture, which is the most important sector in the economy, is exempt from direct income taxes.⁵

Under the Constitution of Pakistan, farmers have been exempt from taxes on their incomes from agriculture and only the Provincial Governments are permitted to levy a land tax.⁶ The Federal Government cannot directly impose taxes on agricultural incomes or land. The political power of large landowners has prevented the Federal Government from seeking

³This was partly due to the lowering of income tax rates in the last quarter of the fiscal year (which starts July 1).

⁴It is estimated that income taxpayers number only one million (out of a population of 135 million).

⁵Agriculture accounts for around 25 percent of GDP, employs one-half of the labor force, and is the largest earner of foreign exchange.

⁶The four provinces of Pakistan are: Balochistan, North West Frontier Province (N.W.F.P.), Punjab, and Sindh.

such a change in the Constitution, as well as constrained the Provincial Governments from utilizing their legal authority to impose taxes on agricultural incomes. At present, the only direct tax on agricultural producers is the provincial land revenue (land tax) whose yield in 1996 was a paltry Rs. 1.3 billion, or 0.4 percent of total tax revenues. The Federal Government has also instituted a small wealth tax on agricultural land, but revenues from this tax are about one-tenth of the land revenue.

Given the tax exemption on income that it enjoys, the agriculture sector has become a legal, and sometimes illegal, tax shelter for other forms of income. To avoid income taxes, transfers from the other sectors of the economy to agriculture are commonplace. Thus, it is estimated that applying the same tax on agricultural incomes as in other sectors would yield substantial revenues over the medium term. Clearly, exempting agriculture from taxation imposes a heavy burden on the rest of the economy as well as a significant loss of revenues for the budget.

As part of the ESAF and EFF programs negotiated with the IMF in September 1997, the Government of Pakistan committed itself to a medium-term strategy to tax agricultural wealth and incomes. This policy is to be implemented in two stages. First, there would be a uniform land-based tax for all provinces by mid-1998, at the level prevailing in the largest province (Punjab).⁷ In the second stage, the Federal Government and the provinces agreed to complete a study on moving from the land-based tax to an agricultural income tax in the medium term. By the terms of the agreement, an agricultural income tax is expected to be in

⁷There is some confusion as to whether this new tax would be in addition to the existing land revenue or replace it. In the major province of Punjab, the land revenue has been abolished.

place some time in the early 21st century. In the meantime, however, little increase in revenues from agriculture is likely to be generated.

The purpose of this paper is to make the case for taxing agricultural producers directly in Pakistan. It first describes the history of failed attempts to underscore the difficulty in introducing what most economists and policymakers would regard as sound tax policy. The second part of the paper examines various alternatives in implementing direct agricultural taxation in Pakistan. Based on the analysis, the paper proposes an agricultural income tax system that could be implemented within the framework of the existing Constitution in the next three to five years.

II. DIRECT TAXES ON AGRICULTURE IN PAKISTAN: SMALL STEPS AND FAILED ATTEMPTS

A. General Aspects

There are two basic issues in agricultural taxation that need to be clearly identified and separated. First, there is the issue of the intersectoral transfer of resources from agriculture to other sectors brought about by explicit or implicit indirect taxes on the former. Most governments have resorted to a variety of indirect explicit taxes as well as implicit taxes (including price controls on output, overvalued exchange rate, etc.) on the agriculture sector, usually at the cost of distorting the allocation of resources and loss of efficiency [Krueger 1996]. Second, there is the issue of taxation of income and wealth of individuals irrespective of residence or which activities generate their income and wealth. In this context, it is

important that a distinction is clearly made between the issue of tax burden on the agriculture sector and the issue of taxes on personal incomes and wealth of landowners and farmers.⁸

Taxes on agricultural commodities have been a major source of government revenue in most developing countries. In spite of the efficiency losses, governments have given preference to this form of taxation for political and administrative reasons. Similarly, they have imposed implicit taxes on agricultural products, but with limited gain to government revenue and gross distortions in resource allocation. In recognition of these costs, since at least the mid-1980s, governments in most developing countries have been moving away from the high level of indirect explicit and implicit taxes on agricultural producers [UN, FAO 1993].

Traditionally, agricultural land has been taxed simply as a natural resource creating rent. In recent times, tax authorities have used three methods to assess the tax on agricultural land: (i) an in rem (property) tax based on land area; (ii) a tax based on the market value or net income of the land; and (iii) a tax based on objective measures which are proxies for land productivity or potential income of land [Skinner 1991]. The first approach has been by far the most common. Only in very few developing countries have governments introduced a tax on actual or presumed agricultural income. The experience is that the revenue yield is often very limited due mainly to problems in the assessment and collection of these taxes. In countries where attempts have been made to transform the in rem land tax (based on land area alone) into in personam land tax (based on land value) the results have been no more

⁸A large number of studies on tax reform in developing countries, using both theoretical models and case studies, have emerged since the late 1980s. See, for example, Newbery and Stern (1987); Ahmad and Stern (1991); and World Bank (1991). While in the literature on agricultural taxation this distinction is clearly made, the discussions in Pakistan have almost always ignored it.

satisfactory. In some countries, taxation of agricultural land as property is left to local (state or municipal) governments, in which case the revenue from this tax has been a reasonably good source of the provincial (state) or municipal income. Land tax is, therefore, closely linked to the services provided in return by state and municipal governments [UN, FAO 1993].

In the context of direct taxes on agricultural producers, a distinction must be made between a tax on income irrespective of the source from which it is derived and a tax on wealth (property) whether mobile or immobile, including agricultural land. Ideally both of these taxes should be a personal liability of the individual and based on actual (earned) income and assessed (market) value of property. In practice, particularly for agricultural income and land, there are severe constraints on assessing actual income and the value of land [Skinner 1991]. As such, a second best approach is to estimate presumed income and value of agricultural land. However, the estimation of presumed income and value of property is also not without serious difficulties. An important aspect of tax policy is that the administration of tax has to be relatively effective--it minimizes tax evasion--and inexpensive--it minimizes rent-seeking.

There is considerable evidence that governments in Pakistan, through a combination of indirect and implicit taxes, have in the past taxed the agriculture sector quite heavily and transferred substantial resources to other sectors in the economy [Pakistan 1993]. These policies have generated some revenues for the government, but in the process have distorted the allocation of resources. Price subsidies on agricultural inputs, though substantial, did not compensate agricultural producers for the losses suffered through price discrimination and export taxes. However, in the last four to five years, considerable changes in the procurement

(support) prices of crops, adjustment in the exchange rate, and liberalization of trade in general have significantly reduced the resource transfer from the agriculture sector [Khan 1991; Khan 1994]. The measures taken in the 1990s have certainly weakened the arguments of the farm lobby in Pakistan that agriculture is already heavily taxed.

In Pakistan, there are only two forms in which the government taxes landowners directly and includes their revenue in its budget: (i) land revenue as a tax on land that generates incomes and (ii) wealth tax on agricultural land as immovable property. The former is assessed and collected by the Provincial Governments and the latter by the Federal Government. In 1982/83 the military government introduced, as part of its policy of Islamizing the economy, the “ushr” levy on Muslim landowners in lieu of land revenue, but the ushr revenue was not to be part of the government’s budget.

B. Land Revenue as a Tax on Land

In 1947, the land revenue systems in the provinces of Punjab, N.W.F.P. and Sindh had significant differences in the structure and methods of assessment.⁹ However, the West Pakistan Land Revenue Act of 1967--which was in fact a copy of the Punjab Land Revenue Act of 1887 enacted by the British colonial administration--introduced a uniform basis for land revenue in what was then West Pakistan, comprising all four provinces of today’s Pakistan. The land revenue rate, determined for each revenue “circle” at the time of “settlement”, is a maximum of 25 per cent of the “net assets”. The net assets are calculated as the value of gross produce minus the “normal” cost of cultivation of crops in the case of owner-cultivators or the land rent received by landowners minus the charge of collecting the rent from tenants.

⁹A detailed study of the land revenue system, with historical background and sources, has been done by Khan (1981, Chapter 8).

The period of each revenue settlement is a minimum of 10 years and maximum of 20 years. Land revenue is paid in cash and is determined on the basis of information about the average crop area matured, crop yield and prices. The revenue rates are classified by soil type for a village or group of villages in the revenue circle. Since the revenue settlements in most districts of West Pakistan were completed before 1947, the government raised the basic revenue rates by 25 percent on an ad hoc basis and adopted the new rates for all of West Pakistan in 1967. Even after the division of West Pakistan into four provinces in 1969, no change was made in the land revenue system until after the splitting of Pakistan into two separate countries in December 1971.

The first change was made in the land revenue system of Sindh after the promulgation of the Land Reform Act of 1972 in Pakistan. A flat rate of land tax was introduced, based on the predetermined value of the Produce Index Unit (PIU), first used in Punjab soon after 1947 to settle the claims of Muslim refugees from East (Indian) Punjab and then extended to all districts of West Pakistan for implementing the Land Reform Act of 1959.¹⁰ The flat rate in Sindh was simply the revenue demand of the three previous years (1969-1972) divided by the PIUs per unit of land in each village (deh) according to the soil type and mode of irrigation. The flat rate was fixed for five years, 1972/73 to 1977/78, after which a new rate was to be established on the basis of revised PIUs per hectare.

Major changes in the land revenue system of all provinces were introduced in November 1975. All irrigated landholdings of up to 5 hectares (or double this size for unirrigated holdings) were exempted from the payment of land revenue, which was a major

¹⁰The PIU is a measure of the productive capacity of agricultural land. It is used to establish the equivalence of parcels of land in different parts of the country.

departure from the existing system in which all landowners had to pay land revenue to the state. This concession was part of the promises made to the “peasants” by the Pakistan People’s Party (PPP) government. All larger size landholdings were required to pay increased rates per hectare. The PPP government followed these changes by the most radical reform in the land revenue system in January 1977: in the Finance (Supplementary) Act of 1977, the Federal Government replaced the land revenue (tax) by a tax on agricultural income, except for owners of holdings of 10 hectares or less. The Finance Act of 1977 represented a breakthrough in that it removed the exemption of agricultural income from taxation. This change in the land revenue system was part of the agricultural reform package--including the Land Reform Act of 1977--offered by the PPP just before the national and provincial elections in March 1977. Under the new system presumed income from agriculture would be determined on the basis of the number of PIUs per hectare, which had not been changed in Pakistan since the late 1940s.

After the coup d’etat in July 1977, the military government suspended the Finance Act of 1977, restored the tax exemption on agricultural income in the Income Tax Ordinance of 1979, and reintroduced the land revenue with new (higher) rates starting from the 1976/77 crop season. Another major change affecting the land revenue system was introduced by the military government in 1982/83 as a result of the Zakat and Ushr Ordinance of 1980. The Sunni Muslim landowners would pay, in lieu of land revenue, the ushr levy at the rate of five per cent of the gross value of output in cash subject to a basic exemption, but the Shia Muslim and non-Muslim landowners would continue to pay the land revenue. The ushr liability on landowners would be assessed and collected by the Local Zakat and Ushr Committees.

After winning the national elections in 1988, the new PPP government made two major changes in the Zakat and Ushr Ordinance of 1980. First, in the Finance Act of 1989, the autonomy of Zakat Councils was reduced and the zakat and ushr funds could be used for a variety of social purposes. Second, the Finance Act of 1990 placed the responsibility of assessment and collection of the ushr levy with the provincial Land Revenue Departments. These changes were intended to increase the government's intervention and expand the ushr base (as a social safety net) for providing relief to the poor in rural and urban areas. In 1991, the Islami Jamhuri Ittehad (IJI) government appointed a committee to examine the means to improve the system and reverse the falling trend in the ushr revenues, but so far no measures to achieve this were implemented in the provinces. In the meantime, the annual ushr collections have fallen from Rs. 260 million in 1983/84 to less than Rs. 50 million in 1995/96. This is considered to be the result of the fact that the administration of ushr has been grossly neglected and that Sunni Muslims have made false declarations about their religious identity to avoid the (higher) ushr levy.

C. Wealth Tax on Agricultural Land

The wealth tax was introduced in Pakistan in 1963, following the recommendations of the Taxation Enquiry Committee of 1959. Under the Wealth Tax Act of 1963, all movable and immovable wealth is subject to tax, with a basic exemption of Rs. one million or a self-occupied house. However, in 1970 the Wealth Tax Act of 1963 was amended to exempt the owners of agricultural land whose income from agriculture was not liable to any tax. This would exclude everyone who owned agricultural land but did not have taxable income from nonagricultural sources, since agricultural income was exempt from taxation. The other problem was that the value of land for wealth tax was set at Rs. 10 per PIU and, since no

individual was allowed to own a land area with more than 8,000 PIUs under the Land Reform Act of 1977, the maximum value of land (Rs. 80,000) would be well under the allowable exemption of Rs. one million.

In 1990, the assessment value of agricultural land for wealth tax was raised from Rs. 10 to Rs. 200 per PIU in two stages by July 1994. However, these changes in the Wealth Tax Act of 1963 made little impact on tax receipts from agricultural land because of at least four factors: (i) the basic exemption of Rs. one million; (ii) the low rate of assessment of the PIU; (iii) the unchanged number of PIUs per hectare since the late 1940s despite the substantial increase in the productive capacity of land due to expanded irrigation, use of modern inputs, development of roads, and urbanization; and (iv) the generous deductions announced by the Federal Government in 1994 after removing the wealth tax exemption for those earning income from agriculture.

In the 1996/1997 budget, the Federal Government raised the assessment rate of land to Rs. 400 per PIU and removed some of the exemptions and loopholes which led to the revenue increase from Rs. 30 million in 1994/95 to Rs. 110 million in 1996/97. However, the problem is that the base of land valuation is still the outdated number of PIUs per hectare. It is obvious that, without the revaluation of land, either by revising the PIUs per hectare or using market value, wealth tax receipts from agricultural land will remain very modest.

D. Current Situation and Proposals

During the 1980s the question of changes in the direct taxes on agriculture, including land revenue, ushr and income tax, was examined by at least three expert committees.¹¹ The

¹¹The three committees were: the National Taxation Reform Commission (1986), National Commission on Agriculture (1988), and the Committee of Experts on Taxation of Agricultural Incomes (1989).

majority view was against introducing a tax on agricultural income. All the committees have emphasized the need to improve the assessment and collection of land revenue and ushr, but without specific recommendations on the methods of implementation. It is well known that the tax exemption for agricultural income has been misused by individuals who receive income from other sources, like industry, trade, etc. Also, taxes are avoided (and evaded) on nonagricultural incomes in several ways. For example, agricultural land would be purchased and then claimed as the source of other incomes; agriculturists would claim income generated from other businesses as agricultural income; and fictitious loans made by agriculturists to industrialists (and traders) help the latter class of taxpayers to avoid tax on their nonagricultural incomes. A minor policy reform with respect to agricultural income was introduced in the Finance Ordinance of 1988 (Ordinance II of 1988). It amended the Income Tax Ordinance of 1979 (Second Schedule in Part I Clause I) to include agricultural income (if any) in the "chargeable income" for determining the tax rate for nonagricultural incomes. This so-called clubbing formula, introduced with the Federal Budget of 1988-89, however, has had no major impact on tax evasion and the income tax revenue.

At the end of 1990, the IJI government appointed a Taxation Committee, which was in favor of a tax on agricultural income, but in its view the Federal Government did not have a clear constitutional authority to introduce a tax on agricultural income. Therefore, the Committee recommended that the Federal Government should redefine the concept of agricultural income by excluding the rental part received by landowners and incomes earned from orchards, livestock and poultry farms. However, no action was taken on this recommendation. The persistently high fiscal deficit and the pressure of the international donor community since the early 1990s have moved up the issue of reforming the land

revenue system on the government's policy agenda in Pakistan. The Caretaker Government, after the dismissal of the Pakistan Muslim League (PML) government in July 1993, took two measures in September-October 1993. First, it amended the Wealth Tax Act of 1963 and removed the exemption on agricultural land as immovable property for wealth tax purposes. In February 1994, the newly-elected PPP government enacted a somewhat watered-down version of the Wealth Tax (Amendment) Ordinance in line with the recommendations of the Prime Minister's Task Force on Agriculture to justify the changes in the original legislation [Pakistan 1993]. Second, it issued ordinances in all provinces to introduce a flat tax rate of Rs. 2 per PIU on all landholdings above 4,000 PIUs to replace the existing land revenue.

The elected provincial governments did not follow up on the provincial tax ordinances on presumed agricultural income (based on existing PIUs) issued by the Caretaker Government in September-October 1993. Subsequently, three provincial legislatures, excluding Punjab, enacted the land tax legislation which introduced a tax on presumptive income at the rate of Rs. 2 per PIU. The Provincial Government in Punjab--the most important agricultural area--took no concrete action to legislate and implement the new tax system based on even the unrevised PIUs.

Finally, in December 1996, at the initiative of the new Caretaker Government, each Provincial Government introduced an "Agricultural Income Tax Ordinance". However, the proposed tax was not on agricultural income, but on either land or crop area under different conditions of irrigation and adjusted for the size of landholding.¹² Since February 1997, the

¹²The Provincial Governments abolished the existing land revenue and replaced it by the so-called agricultural income tax. In other words, the new tax is not in addition to the land revenue. See, for example, the Punjab Land Revenue (Abolition) Ordinance 1997 issued on July 17, 1997.

elected Provincial Governments in Punjab and N.W.F.P. have included a modified form of the land tax in their respective Ordinances. The elected governments in Sindh and Balochistan have legislated a tax system based on the crop area and mode of irrigation. In the Sindh Agricultural Income Tax Ordinance there is stipulation that, starting in July 1998, the tax system will be based on the net income of landowners and leaseholders. However, no details have been provided so far about the proposed tax on net income.

The two legislated tax systems--one in Punjab and the N.W.F.P. based on land area and the other in Sindh and Balochistan based on crop area--are erroneously labeled as "agricultural income tax" and each is flawed. There are at least three problems with both systems. First, they do not take into account the differences between actual and presumed income generated from land. In fact, the 1994 agricultural tax legislation in the provinces included a far better tax regime since it was on presumed income (measured in PIUs per hectare) of the landowner, although the flat tax rate was based on the outdated PIUs. Second, in three provinces (Punjab, Sindh and Balochistan), the exemption for the owners of landholdings of up to 5 (and 6.5) hectares is far too generous, especially in the irrigated areas, considering multiple cropping and crop yields and prices. Finally, at least for the irrigated areas, the tax rates per unit of land/crop are far too low compared to the net value of crop output. It is interesting to note that the provincial governments have made no estimates, based on the enacted tax structures, of the expected agricultural tax revenue for fiscal 1997/98.

Since the passage of the provincial Agricultural Income Tax Ordinances, the farm lobby, particularly in Punjab and Sindh, has agitated a great deal against the new tax system and seems to have succeeded in frustrating the efforts of the Provincial Governments to meet

the revenue target (of around Rs. 1.5 billion) for the fiscal 1996/97.¹³ Three Provincial Governments (in Punjab, Sindh and the N.W.F.P.) have reportedly collected only Rs. 337 million so far. In the meantime, the Federal Government has prepared a schedule of activities for the implementation of a three-year (1997/98 to 1999/2000) economic stabilization and reform program in collaboration with the IMF, the World Bank and other donors. In accordance with the schedule, the Federal and Provincial Governments have decided to appoint a "Technical Committee" to harmonize the agricultural tax regime between provinces and computerize land records for the "refinement" of the existing tax structure. In the short run, the Federal Government seems to have accepted the graduated land tax of Punjab and N.W.F.P. as the model for the other two provinces.

III. IMPLEMENTING DIRECT AGRICULTURAL TAXATION IN PAKISTAN

A. Agricultural Income Tax

At this stage, a second best method may have to be applied in restructuring the direct tax system for farmers and landowners in Pakistan. A global concept of tax on actual income of a person from agriculture cannot be administered efficiently or effectively for several reasons. First, most agricultural production is not organized or conducted on a commercial basis. Second, income and expenditure cannot be verified since records and accounts of cash flows are not kept. Third, tenurial arrangements are complex and not well recorded. Finally, the requirements of tax administration would be far too expensive and beyond the existing

¹³It is well known that the leaders of "Save the Farmers Movement" (Kissan Bachao Tehrik) are political heavyweights and enjoy substantial support both inside and outside the government. It is because of their pressure that the provincial governments in three provinces have revised their Agricultural Income Tax Ordinances.

institutional capacity. Therefore, the idea of taxing verifiable actual personal income from farming may have to wait for some time. However, it would be quite appropriate and practical to apply the existing federal income tax law to incomes earned from commercial orchards and livestock farms, which are reasonably well organized and whose numbers have been increasing in recent years. The owners of these enterprises and their agricultural land should not then be liable for any other tax, except that their land should be subject to the Federal Wealth Tax Act.

There are two important issues with regard to a tax system based on presumed income of a person from agriculture. The first one has to do with the identification of the person whose presumed income has to be the basis of taxation. Agricultural producers (farmers) can earn income, individually or jointly, in four ways: (i) self-cultivation of owned landholding; (ii) cultivation of land leased or rented from someone else on a fixed rent or sharecropping basis; (iii) renting out land to others (sharecropping tenants or lessees); and (iv) nonagricultural professions, services, businesses.

The only verifiable records in Pakistan are on landownership. Even if these records are authentic, there is no way to verify the extent of landownership of an individual since he may own land in different revenue circles. Further, there are either incomplete records or none at all for rental (lease) contracts on agricultural land. While the landownership titles are recorded on an individual (personal) basis, the landholdings are often cultivated or managed jointly within the family. All of this means that the concept of presumed income will have to be applied to the landowners and leaseholders by linking the tax on presumed income to the individual (person) through the ownership and lease of land.

Presumed agricultural income for tax purposes can be defined simply as the annual gross value of output (GVO) minus the cost of production (COP). There are several ways in which the GVO and COP for the landholder (income receiver) can be estimated. The GVO can be calculated by either the actual annual crop output reported by the landowner and verified by the provincial revenue official--by either direct observation or by crop-cutting experiments in the village--or the land revenue official verifies the actual "matured" area of each crop of the landowner and multiplies this area by the average yield on the basis of crop-cutting experiments. The annual output of crops thus reported can then be multiplied by the average price of each crop in the village or in the proximate markets during the post-harvest season.

The next problem is to estimate the COP or the expense incurred in producing the crop output. Here the estimation and verification of the cost of cultivation for each landowner would be almost impossible. A concept of presumed expenses can be used instead. Several farm management studies in Pakistan support the assumption that, on a per hectare basis, COP varies between 35 and 45 per cent of GVO. So the presumed taxable income (PTI) can be estimated for the owner-cultivator as follows:¹⁴

$$PTI = (GVO - 0.40 GVO)$$

There should be no tax liability for the landholders (income receivers) whose PTI is up to the level of tax exemption allowed by the federal income tax rules applied to other sources of income. It would be equivalent to the personal exemption allowed by the Federal Government under its income tax law. For all other persons there could be either one rate of

¹⁴For the landowners leasing out their land and others leasing in land, the tax formula will have to be adjusted to determine their net income.

taxation (say the average tax rate under the federal income tax law) or three graduated rates consistent with the existing tax rates for incomes from other sources. The tax rates, and the basic tax exemption, should be revised as conditions permit or when the rates on nonagricultural incomes change.

B. Agricultural Land Tax

An agricultural land tax in practice is not a pure site value tax on the implicit value of land excluding improvements. There are three options for the taxation of agricultural land:

(i) impose a tax on land area alone such as it now exists in Pakistan and several other developing countries; or (ii) impose a tax based on the market value of land, as has been tried in some Latin American countries; or (iii) impose a tax based on some objective measures, such as the PIU, which can act as a proxy for the potential income or productive capacity of land.

Tax on Agricultural Land Area

The tax on land area (in rem tax) is the easiest to administer. However, it has several serious flaws. For one thing, its revenue potential is limited by the maximum acceptable burden on the least productive land. Also, if for administrative convenience and vertical equity the government exempts a high proportion of landowners with holdings of small size, the tax rate on the rest will have to be very high or steeply graduated. This will quickly lead to changes in land records or fake land transfers on a mass scale. Another disadvantage of the tax based on land area is that the tax burden as a fraction of land value is largest on the least productive land. If the poor farmers own mostly less productive land, then the in rem tax is regressive. Finally, this tax would tend to be inequitable horizontally as well. Productive capacity of agricultural land may vary a great deal between landowners and areas in which

their land is located, depending upon soil quality, availability of water, crop mix, distance to markets, and so on. How can two pieces of land of equal area (say one hectare) be regarded as equivalent for the tax purpose if their productive capacity or potential income differs significantly? In view of these considerations, the recently-legislated land tax (revenue) system in the provinces based entirely on area--with difference only between irrigated and unirrigated lands--has no sound basis for the generation of substantial revenue and horizontal equity in Pakistan.

Tax on the Market or Rental Value of Agricultural Land

In most countries, urban real estate is taxed on the basis of its market (or rental) value to generate revenue for the Local (Municipal) level governments. The same principle has been tried on agricultural land in some Latin American countries. There are several problems involved here. First, if self-assessment of the market value of agricultural land is permitted, without penalties for underassessment, it will generate very limited amount of revenue given the landowners' natural tendency to be modest in assessing the value of their land.

Alternatively, a market-value land taxation could be imposed on each plot of agricultural land using a cadastral survey and reassessing the land value at regular intervals. This can, however, be a very costly exercise, as has been the experience in Brazil [Skinner 1991].

The agricultural land market in most areas of Pakistan is both limited and highly imperfect. The imperfection is directly related to the high concentration of landownership and its influence on the sharecropping tenancy. The market value, if verifiable, will thus reflect the market imperfection and not the income-generating (or productive) capacity of land. Also recorded transactions grossly under-report the market price or value of land and other property. In the rural areas, market transactions for agricultural land are still quite limited and

their records are either incomplete or involve fake transfers of land between relatives. It has also been suggested that the tax on agricultural land could be linked to its rental value, for which information on land rents may be collected through the household and farm surveys. Here again there would be similar conceptual and practical problems as encountered in assessing the market value of agricultural land. Further, rent payments are made in several ways, ranging from the crop shares in kind and cash to the fixed rents in kind and cash.

Tax on the Productive Capacity of Agricultural Land

A practical alternative to the tax on actual or presumed income would be to tax the landowners on the basis of productive capacity (or potential income) of their agricultural land. This option will accommodate changes in land productivity and the tax-paying capacity of the individual landowner or landholder, avoid the need for market assessments, and be adjusted for inflation by increasing the tax rate on the (constant) quality index of land. In the context of Pakistan, it would be appropriate to use the PIU as an (objective) measure of the productive capacity of agricultural land, since it takes into account the conditions of soils and water and the gross value of output. As stated earlier, the PIU was designed to establish the equivalence between agricultural land parcels located in different parts of the country. Since then the PIU has been used as the basis for: (i) the ceiling on landholdings and land titles under the Land Reform Acts of 1959, 1972 and 1977; (ii) agricultural loans from commercial banks and the Agricultural Development Bank of Pakistan; (iii) the federal wealth tax on agricultural land as immobile property; and (d) the flat rate of land tax (revenue) in Sindh in the 1970s.

The problem in using the PIU as a proxy for productive capacity of land or its potential income is that it has not been revised since the late 1940s and early 1950s when most

of the land claims of Muslim refugees were settled. Successive governments have made public pronouncements and promises after the land reform of 1972 to revise the PIU in accordance with changes in the soil and water conditions, crop mix, etc., but so far nothing has been done. No one has questioned seriously the notion of PIU as a reasonable proxy for the productive capacity or potential income of land. However, it has been argued that, given the data requirements, it would be very expensive and time consuming to revise the PIU in the provinces. The task is not as difficult as is sometimes argued, since the information and resource requirements are in fact within the capacity of (almost) all of the provincial Land Revenue Departments.

The revised number of PIUs per hectare in each revenue circle should form the basis of taxation of landowners or their landholdings. The basic tax exemption in terms of PIUs can be set at a level that excludes a vast majority of small landowners and landholdings with a small number of PIUs as in many unirrigated and waterlogged areas. A proportional rate of tax per PIU should be used for all landholdings or landowners above the exemption level. A graduated tax rate will create serious problems of tax evasion and may lead to much rent-seeking. The basic exemption for most landowners would give progressivity to the tax regime.

Using the data on ownership holdings and estimates of revised PIUs per hectare based on some restrictive assumptions,¹⁵ three reasonably practical revenue scenarios show that the Provincial Governments of Punjab, Sindh and N.W.F.P. can receive significant additional income from the landowners by a tax system based on the productive capacity or potential income of agricultural land.

¹⁵A simple yet practical method for updating the PIUs is given in Appendix I.

In the first scenario, the basic exemption could be set for the landholdings of up to two hectares with a tax rate of Rs. 1 per PIU for all other landholdings. This can generate Rs. 7.3 billion in tax revenue paid by owners of about two million landholdings at Rs. 443 per hectare or 11 percent of their net crop income of Rs. 4,000. The second scenario is that the basic exemption is raised to landholdings of up to five hectares, while keeping the tax rate at Rs. 1 per PIU. In this case, the revenue yield would be Rs. 5.3 billion paid by owners of nearly 800,000 landholdings with the average tax burden of Rs. 506 per hectare or 13 percent of their net crop income. The third option is to impose a tax rate of Rs. 2 per PIU for the landholdings of over five hectares, which can potentially generate Rs. 10.6 billion in provincial revenue. Here the tax burden would be Rs. 1,000 per hectare or 25 percent of the owners' net crop income.

It is important to point out that the collected revenue is likely to be smaller than the estimated potential revenue in each of the three scenarios because of leakages due to incomplete assessment, poor administration and rent-seeking. Even with the assumption of a shortfall of 25 percent, the collected revenue would represent a significant addition to the provincial income. For example, the ratios of the tax based on revised PIUs to the total provincial taxes and revenue would be in the range of 28 to 43 percent and 17 to 29 percent, respectively. At present, the ratios of land revenue to the provincial taxes and total revenue are only 11 and 6 percent. The contribution of the proposed tax to direct taxes collected by the Federal and Provincial Governments would be in the range of 5 to 10 percent, whereas it is now less than 2 percent. The overall tax burden on agricultural producers would still be reasonably low. The ratio of tax revenue to the crop value-added (national accounts) would be in the range of 1.4 and 2.7 percent, compared to the current ratio of 0.45 percent. The

revenue paid per hectare (of cultivated land) would be between Rs. 185 and Rs. 369 per hectare as compared to the current revenue of Rs. 60 per hectare.

IV. CONCLUSIONS

The debate on agricultural taxation in Pakistan has taken two related forms. For a long time the discussion centered on whether agricultural incomes and wealth should be taxed at all. When it became generally accepted that there was no compelling economic reason for exempting agricultural incomes, the issue turned to how should agricultural producers be taxed. In the Pakistan environment it is difficult to tax agriculture, and thus administrative constraints have become the new rationale for excluding agriculture from the tax net.

This paper suggests a way that the Government of Pakistan could implement a system of agricultural taxation within the existing Constitutional framework. In the short term, say over the next two to three years, the Provincial Governments should revise the number of PIUs per hectare, since they are completely out of date in terms of reflecting market values or productive capacity of farm lands. A simple method of making such revisions is outlined in the paper. An adjustment of PIUs alone would yield significant revenues in the short run. In the longer run, possibly from three to five years, the Provincial Governments could move to taxing net agricultural income, using the productive capacity of land or presumed income as the base.

At the technical level, there seem to be good arguments for the type of agricultural tax system suggested here. However, two important considerations apply. The first is a political economy issue, given the power of the landlord lobby. Even if the Federal Government was convinced of the need to tax agricultural producers, it would need to persuade provinces to

enact such a tax. In fact, there are levers that the Federal Government can use towards this end. The Provincial Governments depend on federal tax revenue transfers, grants-in-aid, and foreign aid. Linking such transfers to the capacity of Provincial Governments to generate tax revenues from agriculture may well offset the political pressures of the landowners.

Second, there is the practical issue of the administrative capacity of the provincial revenue administration.¹⁶ According to recent studies of the land revenue system in Pakistan, it is clear that much work needs to be done to update the land title records. Also, digitized soil maps are required to establish land classification and values by soil, water, and crop conditions. These are necessary requirements to estimate land values and ownership, and thus presumed incomes for tax purposes, and it will take time before accurate estimates are available. Therefore, a period of three to five years may be needed to implement a full-fledged agricultural income tax.

In conclusion, it is now generally accepted in Pakistan, even by the landowning groups, that agricultural producers should pay their fair share of taxes. While equity arguments are important, the need to address the severe fiscal problems the country is facing has led to a fundamental rethinking on agricultural taxation. Revenues from a broad-based agricultural income and wealth tax would help considerably in achieving fiscal balance. How to design and implement the tax is a difficult question, but not an impossible one to resolve. As argued in this paper, technical solutions are possible, but ultimately taxing agriculture in Pakistan will require considerable political will, which unfortunately has been lacking except in a few isolated instances of the country's history.

¹⁶The last major review of the provincial revenue administrations was undertaken by the military government in 1978. See Khan (1981).

REFERENCES

- Ahmad, Ehtisham and Nicholas Stern, 1991, *The Theory and Practice of Tax Reform in Developing Countries* (Cambridge: Cambridge University Press).
- Khan, Mahmood H., 1981, *Underdevelopment and Agrarian Structure in Pakistan* (Boulder: Westview Press).
- _____, 1991, "Resource Mobilization from Agriculture in Pakistan," *Pakistan Development Review*, Vol. 30, No. 4, Part I, pp.457-483.
- _____, 1994. "The Structural Adjustment Process and Agricultural Change in Pakistan in the 1980s and 1990s," *Pakistan Development Review*, Vol. 33, No. 4, Part I, pp.533-591.
- _____, 1997. "Structural Adjustment, Economic Performance and Government Budgets in Pakistan," *Canadian Journal of Development Studies*, Vol. 18, No. 2, pp.279-302.
- Khan, Mohsin S., 1992, "Macroeconomic Balances," in *Foundations of Pakistan's Political Economy*, ed. by W. E Jones and S. Roy. (New Delhi: Sage).
- Krueger, Anne O., 1996, "Political Economy of Agricultural Policy," *Public Choice*, Vol. 87, Nos. 1-2, pp.163-175.
- Newbery, David and Nicholas Stern, eds., 1987,. *The Theory of Taxation for Developing Countries* (New York: Oxford University Press).
- Pakistan, Ministry of Finance, 1993, *Report of the Prime Minister's Task Force on Agriculture* (Islamabad).
- Skinner, Jonathan, 1991, "Prospects for Agricultural Land Taxation in Developing Countries," *World Bank Economic Review*, Vol. 5, No. 3, pp.493-511.
- Strasma, John and Others, 1987, *Impact of Agricultural Land Revenue Systems on Agricultural Land Use* (Burlington: Associates in Rural Development).
- United Nations, Food & Agriculture Organization, 1993, *Agricultural Taxation in Developing Countries: An Overview of Selected Case Studies* (Rome).
- World Bank, 1991, *Lessons of Tax Reform* (Washington).

A Simple Method for Revising the Produce Index Units (PIUs) in Pakistan

The measurement of Produce Index Units (PIU) is based on the gross value of crop output (GVCO) of a unit of land by soil type. Any two hectares of land located in different revenue circles are assigned the same PIU if they are capable of producing an equal amount of GVCO in a year. The PIUs in Pakistan were calculated initially in 1947 and were adjusted as part of the Land Reform Act of 1959. The current values of the PIUs can be approximated by scaling up the original PIUs by the ratio of average GVCO in the current period (PIU_c) and base period (PIU_b) as follows:

$$PIU_c = PIU_b \times [\text{Ratio of Average GVCO in the Current and Base Periods}]$$

The specific steps involved in the calculation are:

- Step 1. Select a "standard" revenue circle in each defined land area (Tehsil/Taluka). Assume the PIUs for this revenue circle were estimated at 64 per hectare in 1959/60 (used by the Land Reform Act of 1959).
- Step 2. Estimate the GVCO per hectare for the base and current periods. These values could be based on the three-year average of the matured area, output and price of each crop. Say, for example, the estimated GVCOs per hectare for the two periods in the standard revenue circle are:
- | | |
|--------------------------|------------|
| Base Period (1957-59) | Rs. 1,500 |
| Current Period (1995-97) | Rs. 10,000 |
- Step 3. In this case, the GVCO has increased by a multiple of 6.67. So the number of PIUs for the standard revenue circle in the current period are:
 $64 \times 6.7 = 427$ per hectare.
- Step 4. Assume that in another revenue circle (within the same Tehsil/Taluka) the GVCO for the current period (1995-97) is estimated at Rs. 8,000 per hectare. The ratio of GVCO of this circle to the GVCO of the standard revenue circle is 0.80 in the current period. Multiply the updated number of PIUs for the standard circle (427) by this ratio to get the updated PIUs for the other circle:
 $427 \times 0.8 = 342$ PIUs per hectare.

Step 4 should be repeated for all revenue circles in each Tehsil/Taluka of each province.

An alternative method would be to repeat the steps used for the standard revenue circle in all other revenue circles in each Tehsil/Taluka. This method would be more accurate, but would require far more time, resources, and data.

Table 1. Changes in the Provincial Land Revenue Collections, 1948/49-1996/97
(In millions of Rupees)

Year	Land Revenue	Provincial Revenue	Provincial Taxes	Crop Value Added	Cultivated Area (millions of hectares)
1948/49	48	317	106	--	15.16
1950/51	47	383	145	--	15.15
1954/55	58	494	202	--	15.31
1960/61	135	844	248	5,011	18.12
1964/65	154	1,053	374	7,066	18.72
1969/70	172	1,325	539	11,102	19.23
1974/75	194	2,024	1,458	23,271	19.55
1979/80	257	2,368	1,809	43,993	20.23
1984/85	253	4,298	3,297	80,126	20.61
1989/90	693	6,619	5,431	115,065	20.94
1994/95	1,095	21,220	12,076	249,991	21.55
1995/96	1,266	20,794	11,665	291,706	21.54
1996/97	—	23,720	14,279	311,195	21.54

Note: These data are from the annual issues of four publications of Government of Pakistan: *Pakistan Economic Survey*, *Statistical Yearbook*, *Land and Crop Statistics*, and *Agricultural Statistics of Pakistan*.

Table 2. Impact of Alternative Tax Structures Based on Revised PIUs¹
(In millions of Rupees)

Scenario	Land Tax Revenue	Provincial Taxes	Provincial Revenue	Provincial Land Tax Revenue:	
				Percent of Crop Value	Per Hectare of Cultivated Land
Actual (1995/96)	1,300	11,665	20,794	0.45	Rs. 60.35
Scenario One	5,475	15,840	24,969	1.88	Rs. 254.18
Scenario Two	3,975	14,340	23,469	1.36	Rs. 184.54
Scenario Three	7,950	18,315	27,444	2.73	Rs. 369.08

¹Figures in the first row are the reported official data for the 1995/96 fiscal year as shown in [Table 1](#). The crop value for 1995/96 was Rs. 291,706 million and the cultivated area was 21.54 million hectares.

Chart 1 Pakistan

Fiscal Deficit, 1992-1996 (In percent of GDP)

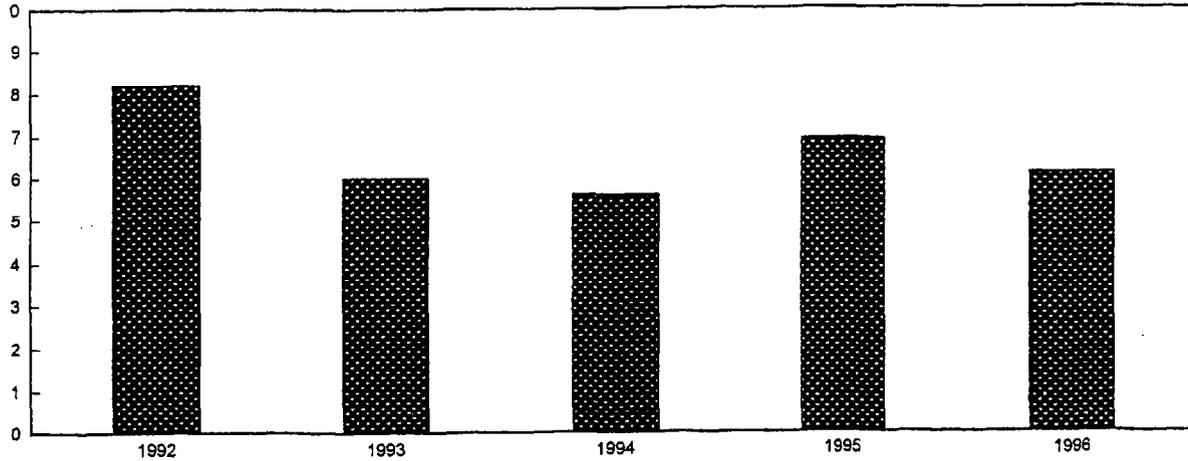


Chart 2

Fiscal Developments, 1992-1996 (In percent of GDP)

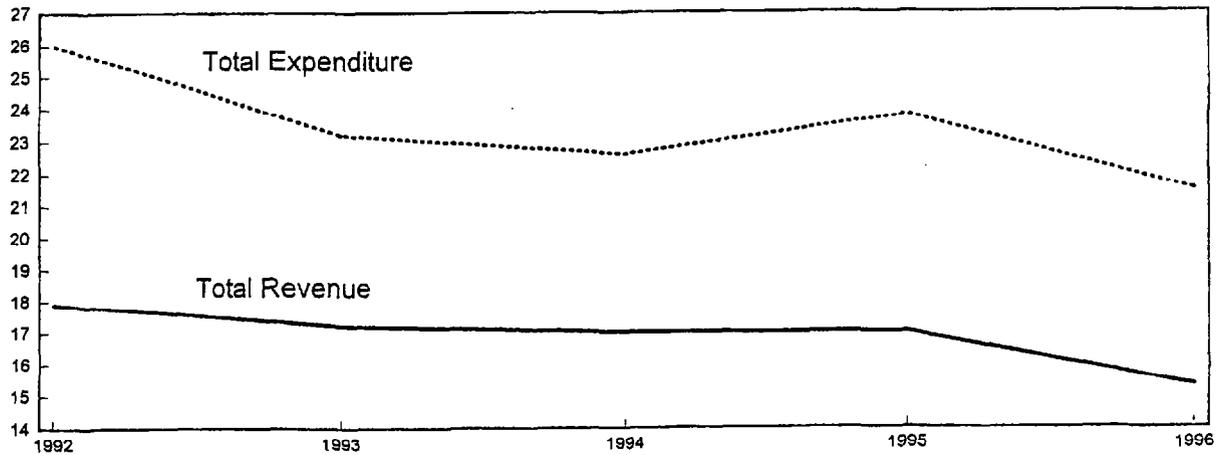
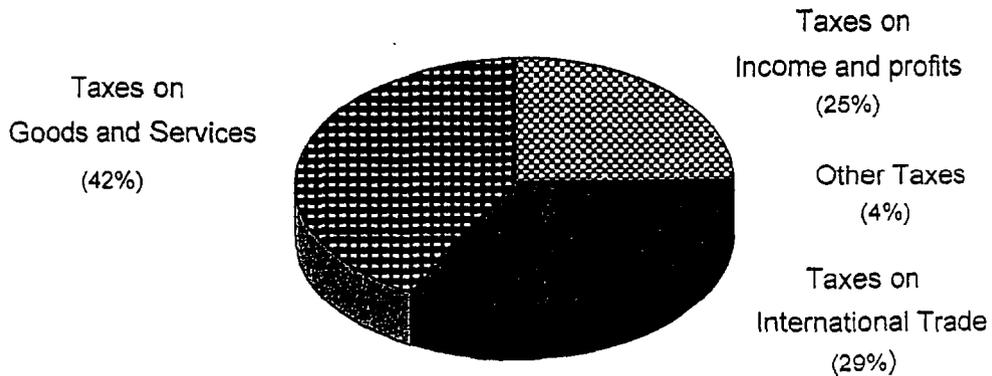


Chart 3

Composition of Tax Revenues, 1995-96 (In percent)





ID_Doc: PPAA/98/4

JOBNO 46318

SERIES PPAA

Docdate 05/01/1998

IMF Paper on Policy Analysis and Assessment

© 1998 International Monetary Fund

This is a *Paper on Policy Analysis and Assessment* and the author(s) would welcome any comments on the present text. Citations should refer to a *Paper on Policy Analysis and Assessment of the International Monetary Fund*. The views expressed are those of the author(s) and do not necessarily represent those of the Fund.

PPAA/98/4

INTERNATIONAL MONETARY FUND

Monetary and Exchange Affairs Department

The Payment System and Monetary Policy

Prepared by Omotunde E.G. Johnson¹

May 1998

Abstract

Achieving the primary objective of price stability without unduly compromising the operational efficiency of the payment system constitutes a major problem for central banks. Routine monetary policy presumes a given institutional and technological framework, including aspects of the payment system. Such a monetary policy concerns itself with intraday and interday credit for payments settlements and with float. Liquidity shocks and panics sometimes pose an additional challenge. In recent years, major and rapid institutional and technological changes in the payment system (mainly to lower risks and augment operational efficiency) have affected the monetary policy decision-making process, particularly in the short run.

JEL Classification Numbers: E51, E52, E58

Author's E-Mail Address: OJohnson@imf.org

¹Prepared initially for Payment and Settlement Systems Workshop at the Bank of England (January 12–16, 1998). I would like to thank William Alexander, Daniel Hardy, Timothy Lane, Nicholas Roberts, Gabriel Sensenbrenner, and Valeria Fichera, for comments on earlier drafts. The views expressed in this paper are those of the author and do not necessarily reflect those of the IMF.



I. INTRODUCTION

Monetary policy considerations happen to be only one factor determining the role that a central bank can play in the payment system,² if we assume, as seems reasonable, that the central bank has, legally, an important role in the conduct of monetary policy as well as in fostering stability and soundness of the financial system. Increasingly, central banks have come to focus on price stability as their primary objective, although some central banks may see maintaining the external value of the currency (the exchange rate) as the primary objective. In any event, this means that the effective conduct of monetary policy has become the primary preoccupation of central banks.

In general, the instruments and the intermediate targets used by central banks to target inflation could be one or more interest rates and some monetary aggregates, such as base money, M1, or M2. In monetary operations, a concerted shift from direct to indirect instruments is occurring world-wide. Hence one observes a decline in use of instruments such as selective credit controls and bank-by-bank credit ceilings, and a relative increase in use of instruments that are market-oriented, where "price" is the basic determinant of allocation, and with access criteria that are transparent, objective, and uniformly applied, where quantitative ceilings get used in conjunction. The latter instruments include reserve requirements, a variety of refinance standing facilities like overdrafts or discount window lending (used at the initiative of the commercial banks), money market refinance instruments (typically in the form of repos and reverse repos and used at the initiative of the central bank), and open market

²For an introduction to payment systems and recent payment system policy initiatives, see Summers (1994), as well as Johnson, et al (1998).

operations (typically in the form of central bank outright purchases and sales of government securities).³

The above developments have occurred in the context, *inter alia*, of a growing complexity and efficiency of financial markets and financial instruments, globalization and integration of markets, and sustained growth in private capital flows. Short-term (daily, weekly, monthly) monetary operations themselves are increasingly conducted with use of a liquidity management framework, which rely on models from basic monetary theory (money demand, money multiplier, etc.), some short-term information system,⁴ and lots of judgement about the short-term evolution of particular parameters and market reactions (especially over the short-term) to policy initiatives.⁵

³In this regard, see, for example, the collection of papers in Baliño and Zamalloa (1997).

⁴Typically a short-term information system will comprise two broad categories of information: one that describes monetary conditions and another set often referred to as supplementary indicators. The first set will include data such as the components of base money, total base money, and broader monetary aggregates. The second category will include data on interest rates, exchange rates, aggregate business activity, aggregate demand, balance of payments developments, government finance, and labor market conditions (wages, employment, unemployment). The more direct is the impact of monetary and credit policy actions by the central bank on a supplementary indicator the more useful is that indicator. Hence, central banks give pride of place to information on interest rates, exchange rates, and other variables that affect or reflect supply and demand conditions in financial markets.

⁵For example, in the very short term, a central bank may want to: (i) develop very detailed procedures for making judgements on the desired change in base money (net of changes in required reserves) that it wants to effect; (ii) make projections of changes in base money likely to originate from so-called autonomous factors (e.g., credit to government, foreign exchange intervention, balance of payment developments, etc., depending on the system) and then (iii) to design its monetary operations, carefully assigning different instruments the roles (contributions) that they should play in bringing about the changes desired in base money.

In none of this does one normally seem to mention explicitly the payment system. In addition, the worldwide growing (operational) independence, and the narrowing of the focus, of central banks, and the resulting preoccupation with inflation as the primary objective of these institutions, have resulted in increasing support for central banks worrying less and less about the effect of monetary policy and operations on the state of unemployment, the level of economic growth, and bank failures. Why in this context should the payment system be treated differently? Here we run across the first concern of the monetary authorities: a moral hazard problem and the challenge of not allowing concerns for the payment system to compromise achievement of monetary objective(s). In other words, the inability to settle by one or more participants could force the central bank to extend special credits to prevent a systemic crisis in the financial system. The dilemma here is greater in the case of the payment system than for the other areas mentioned above, because in the payment system one is dealing with money—typically fiat money—the medium of exchange and the ultimate means of settling debts. In addition, central bank money normally has a monopoly as legal tender and forms the base for all other components of money.⁶

A major payment system initiative in recent years with explicit concern for monetary policy has been the TARGET (Trans-European Automated Real-Time Gross Settlement Express Transfer) System. In November 1994, the European Monetary Institute (EMI) released a note (EMI, 1994) that described why the European Union central banks planned to

⁶As central banks come to see themselves as guarantors against systemic failures in the financial (including payment) system, they have been putting in place risk control measures to reduce systemic risks and hence the probability that they would be called upon to rescue the system from systemic collapse, and, ipso facto, reduce the potential credit risks to the central banks themselves (see also Section III below).

link the national RTGS systems, which were operating (or were about to operate), in line with one of ten principles (principle four) aimed at establishing minimum common features for systems, in the context of harmonization of payment systems in the European Union countries.⁷ A report on the TARGET system (EMI, 1995) gave a detailed description of the future system, explained how that system would be organized, how it would operate, and its possible future links with other payment systems.

The analysis of the existing large-value payment systems in the member countries showed that more than 25 systems were dealing, exclusively or in part, with large-value payments, and that those systems generally were independent and not linked. Indeed, the exchange of large-value payments between countries was relying on correspondent banking arrangements, which, after all, had been assessed to be inconsistent with the requirements for implementing a single monetary policy. Therefore, in line with principle four, the central banks decided that the future European large-value payment system should allow the exchange in real time, on a gross basis, of payments in central bank money, based on the linkage of the RTGSs that operated (or would soon operate) in European Union countries.

⁷See EMI (1993). Principle four stated that each member state should have, as soon as feasible, an RTGS system through which as many large-value and time-critical payments as possible should be channeled. A *gross settlement system* is a transfer system in which settlement of funds or securities transfers occurs individually on an order-by-order basis (that is, without netting debits against credits) according to rules and procedures of the system. In a *real time gross settlement system*, processing and settlement take place in real time (continuously). Since RTGS precludes accumulation of unsettled balances and the associated extension of interbank intraday credit, it helps avoid the systemic disruptions that could result if a major segment of the system could not settle at some deferred prescribed time (typically end of day). This has been a major reason for the popularity of RTGS systems among large value systems owned by central banks. For a more complete discussion of the advantages and disadvantages of the typical RTGS and the typical netting system, see Summers (1994) and Johnson, et al (1998).

Apart from this interesting case, recently, major changes have been occurring in the payment system—because of reforms to address speed, reliability, and financial risks, in the systems, or because of endogenous changes, as financial institutions make payments innovations in search of greater profits—enough to force the monetary authorities to take note of the payment system in the monetary policy decision-making process. This is because of the impact of these changes on money demand and money supply, on the effectiveness of monetary policy instruments, and on the usefulness of certain intermediate targets. Indeed, it is the importance and rapidity of these developments in recent years that have brought payment systems into the normal preoccupation of persons narrowly concerned with monetary policy.

If one turns the tables around and looks at the problem from the perspective of the payment system, one also begins to see why persons who would consider themselves payment system experts have come to take an interest in the relationship between the payment system and monetary policy. The point is that what monetary authorities consider appropriate monetary operations and the stance they take in the conduct of monetary policy can have implications for the operational efficiency of the payment system, because of the effect on the cost and quantity of liquidity. So can monetary policy afford to ignore the operational efficiency of the payment system? If not how much should it care?

In all these areas we still need a lot of theoretical and empirical research work. In the meanwhile, practical people have found various answers. Hence what we have now is a menu of “best practices” that have been shown to foster the achievement of desirable objectives for payment system risk and efficiency as well as for monetary policy, while taking into account

the particular circumstances (institutions, state of economic development, and evolutionary path of the payment system) of countries.

The rest of the paper is divided into three basic sections. Section II deals with issues of routine monetary operations and the payment system—when institutional and technological changes are given—namely, central bank credit policy without sacrificing the main objective of monetary policy and addressing payment system float. Section III discusses liquidity shocks and panics originating from, or aggravated by, the payment system as a major challenge to routine monetary policy. Finally, Section IV discusses the importance of institutional and technological changes for monetary policy and operations, particularly in the short run.

II. ROUTINE MONETARY OPERATIONS AND THE PAYMENT SYSTEM

If the primary objective of monetary policy is price stability, then monetary operations vis-a-vis the payment system should not, under normal circumstances, create unforeseen and unpredictable difficulties for the achievement of this objective. This, we believe, is the most important underlying principle in the analysis of monetary policy and the payment system. A typical payment system expert, we also believe, would want, in addition, that monetary policy not threaten the operational efficiency of the payment system. Coming up with clear principles to reconcile these two objectives remains, in our estimation, a major analytical problem.

When the monetary authorities supply base money to the economic system, in the context of its liquidity management, either to influence the price level directly or to do so via some intermediate target, it presumes a certain institutional and technological framework that in turn affects money demand function, the money supply process (money multiplier), and the

nature of the transmission process relating base money and inflation. The payment system is part of that institutional and technological framework. This would seem a fundamental analytical point to realize in any theoretical and empirical work on monetary policy and the payment system. In general, the most important payment system related institutional and technological factors include: the type of large value transfer system (including key features); the state of development of money markets, especially intraday and overnight markets; the major payment instruments in use; and the reserve maintenance system—in particular whether or not reserve averaging occurs and the extent to which required reserves can be used for intraday settlements.⁸

Between countries, of course, in so far as payment systems differ in important details, there will be differences in real demand for bank reserves.⁹ Hence, other things being equal, if, for instance, one country has a large value transfer system with (daily) net settlement (and

⁸One often-cited disadvantage of a real time gross settlement (large value) system is the risk of gridlock from insufficient intraday liquidity (in terms of clearing or reserve balances) to ensure high operational efficiency of the system. This disadvantage has been addressed, by central banks in various countries, through some combination of sophisticated queuing mechanisms, central bank intraday credit facilities, and central bank reserve management policy (including required reserve policy).

⁹In fact, these differences may be between groups of countries, as there may be strong international linkages between financial change and development of countries conditional on their income levels, their financial market orientation, and their general economic interrelations with each other (say via integration of financial markets or international trade and payments). Thus, in discussing initiatives and developments in payment systems, it has been possible to group countries as industrial countries, middle-income countries, low-income countries, and economies in transition [see, for example, Johnson, et al. (1998)]. Also, at least for the industrial countries, there is evidence of a strong similarity in the behavior of velocity across countries for which a significant part of the explanation is the common institutional and technological changes in the financial systems [see, e.g., Bordo, et al (1997)].

“reasonable” allowance for credit exposures and bilateral debits), while another has a strict real time gross settlement system (with little or no intraday credit from the central bank), the optimal bank reserves per unit of real output to achieve any given rate of inflation is expected to be higher for the latter than for the former country. In so far as reserves comprise domestically generated fiat money—that is they are liquid liabilities of countries’ own monetary authorities—this difference in optimal reserve “need” should not have significant welfare effect and, ipso facto, no particular impact on the ability of the central bank to achieve its inflationary objectives.

Other institutional and technological factors in the payment system also cause differences in optimal reserves-GDP ratios to achieve any given rate of inflation. These factors include: the average length of time taken in processing and settling payments; the account structure of banks, in particular the extent to which these are centralized in, say, a single reserve account at the central bank, and related to that the access to information on unsettled accounts; and the liquidity management expertise in the financial institutions of countries, due both to back-office procedures and to money management activities. Forecasting of system-wide demand for bank reserves (long-run and short-run) both within and across countries is clearly aided by an understanding of these and other factors in the payment system.

A. Central Bank Credit Policy

An important component of monetary policy is central bank refinance policy. Thus, credit policy in the payment system has become a prominent topic. But the discussion has not been confined strictly to monetary policy issues; for those concerned with credit to facilitate the operational efficiency of the payment system, also salient has been the question of how to

limit the credit risk of the central bank. It would seem useful to keep a clear focus on the reasons (risk reduction vs inflation control) for specific procedures and rules in this area, even though the central bank may be the policy maker in both cases. As regards the credit itself, for real time gross settlement (RTGS) systems, the basic issue has been whether the central bank should grant intraday credit at all, while for net settlement systems the debate has centered around the automaticity of overnight credit of the central bank to facilitate settlement.

Interday credit

One could safely say that a consensus has emerged that there should be nothing automatic about lending by the central bank to facilitate end-of-day settlement in payment systems (net or gross). The general view would be that lending for this purpose should be treated as normal refinance credit. In this light, central banks would typically want banks, for end-of-day payment settlement purposes, to use standing facilities of the central bank; would limit banks' access to such facilities, particularly when the credit is uncollateralized; and would charge interest in line with the current stance of monetary policy.

Some central banks will not even lend without collateral. When done for reasons of credit risk, the logic of strictly collateralized lending is clear; it is not so clear that this restriction is necessary for monetary policy reasons. In sum, one could see some logic in limiting the aggregate of such lending, for monetary policy reasons, while asking for collateral, for risk control reasons.¹⁰

¹⁰Of course the demarcation should not be exaggerated, since the requirement of collateral may also have the indirect effect of limiting credit demand.

The relationship between the applicable interest rate(s) on the facilities and market interest rates also varies across countries. Whether or not the lending is collateralized, typically, *changes* in the ratio of the basic interest rate(s) levied on such “refinancing” to current short-term market interest rates are intended to signal changes in the stance of monetary policy. But the normal *level* of such a central bank rate in relation to current short-term market interest rates would clearly depend on whether the credit is seen as facilitating operational efficiency of the payment system, as well as on the normal place that such a rate plays in the structure of central bank interest rates. In this light, short-term overnight rates of central banks can vary all along the spectrum from below market rates (which is rare) to what are essentially penalty rates and an upper bound for market rates.

From a strictly monetary policy point of view, the case appears strong for charging penalty rates on central bank facilities used by banks to complete their (end-of-day) payment settlements. For once the monetary authorities have provided “adequate” overall liquidity to the system, the idea would be to create an incentive for a bank to look to the market first for funds. In that case, when the bank comes to the central bank for (settlement) credit, the central bank receives information about the “tightness” of liquidity in the system or the creditworthiness of the bank. If the nature of the information is clear, the monetary policy decision-making process benefits; the central bank has one more piece of data to help it assess the appropriateness of its monetary policy stance (and also the state of individual banks). Obviously, where financial markets are not active or integrated nationally, central banks may need to play a positive role in developing such markets and, in the meanwhile, more actively engage themselves in normal financial market (funds recycling) activities than otherwise.

Intraday credit

From a monetary policy perspective, the most interesting, even if obvious, questions related to intraday credit in the payment system, are: why is there a demand for intraday credit and what factors determine this demand? who should supply intraday credit—the central bank, the market, or both? does the level of such credit ever matter for inflation? The first question is the one most researched and also, perhaps, least contentious, the second the most contentious, and the third the least researched and yet the most important for monetary policy.

Basically the demand for intraday credit arises especially in real time gross settlement (RTGS) systems.¹¹ For the value of fund transfers that occur during any single day is typically several times the underlying bank reserves available for final settlement. Averaging of reserve holdings for purposes of meeting required reserves and permitting intraday use of all reserve balances for payment settlement purposes helps in satisfying some of the demand for intraday reserves. But this may still not suffice to ensure tolerable operational efficiency of the system. In any event, bank reserves must turn over several times during the day (hence the notion of “turnover ratio”) when settlement is real-time gross.¹² The challenge to policy makers is preventing *gridlock*, avoiding major constraints on payments processing speed and volume

¹¹We are talking here, of course, of explicit credit. In net settlement systems, for instance, banks effectively create inside money by having bilateral debit and credit balances during the day. As others have pointed out before, these would show up clearly as explicit credit if banks, for example, had to borrow in some market to make the payments, rather than accumulate debit balances.

¹²The rapidity with which this turnover ratio increased in the United States can, for instance, be seen from the evidence that the ratio of average daily payments through the major payment networks (for both wire transfers and checks) to average daily reserve balances maintained with the Federal Reserve Banks rose from 0.9 in 1960 to 30 in 1985; see Mengle, Humphrey, and Summers (1987).

that could result in transaction levels well below what the particular payment system can handle.

Part of the institutional and technological environment with which monetary policy must contend in the payment system is, of course, the set of devices available to raise, efficiently, the turnover ratio, and thus support greater gross transfers for any given level of reserves. In particular *queuing* systems get devised. For instance, systems and criteria are put in place to delay payments (“delayed sends”) until covering funds become available. This sort of queuing of payments could be centralized—by having the payments organization or system designing and doing the queuing, as in the SIC system¹³—or it could be decentralized by having the sending financial firms do the queuing.¹⁴

Because of its relatively high demand for bank reserves and the need to take into account considerations relating to monetary policy as well as operational efficiency and credit risk of the central bank, the provision of intraday liquidity is given serious consideration when central banks design RTGS systems. Not surprisingly, central banks come out with different solutions. For example, in their respective RTGS systems, the central banks of China, Germany, Japan, Korea, and Switzerland do not provide any intraday liquidity; the central banks of Denmark, Italy, Netherlands, Portugal, Spain, Sweden, and Thailand provide it through collateralized overdrafts; and the central banks of Australia, France, and the United Kingdom provide it through intraday repos. The central banks of United States (to prime

¹³see “Payment Systems in Switzerland” in BIS (1993), pp. 351-383

¹⁴Some form of queuing may be the required trade-off for the removal of systemic risk under an RTGS system, when the central bank is not prepared to take on the credit risk of substantial intraday lending to banks.

banks within limits) and Mexico provide intraday uncollateralized overdrafts. Also, some RTGS systems (e.g., those of Australia, France, and Hong Kong SAR) have both queuing and intraday credit. When no intraday liquidity is provided by the central bank, the RTGS systems, except in Japan, have sophisticated queuing mechanisms. In addition, Japan and Switzerland, for example, have intraday financial markets, again an important part of the institutional and technological environment.

This brings up another issue of major practical importance, namely, finding indicators to use as benchmarks for the pricing (supply price) of intraday credit, when the central bank provides such credit. As we have argued, a central bank in extending such credit could have both monetary policy objectives and payment system operational efficiency to worry about.

Leaving aside any relevant adjustment for policy signaling, in principle, the price (interest charge) would be equivalent to the sum of the administrative costs, the opportunity cost of the funds, and some adjustment for risk. If a market existed and all risks were internalized, this risk adjustment would include the credit risk connected with the overdraft. In the absence of an actual intraday private market for funds, such as the intraday market in Japan, a central bank must look for proxies. A reasonable approach would start with the interest rate prevailing in a fairly similar funds market and adjust it for risk, maturity, and implicit charges for other services jointly supplied.¹⁵ Another approach would start with some

¹⁵For instance, in the U.S. the rate on day loans used by broker/dealers to finance securities purchases prior to delivery and payment by customers was thought by some as a relevant rate to use to approximate a daylight overdraft price. Naturally, it was recognized that this rate should be adjusted for the factors mentioned above; see Mingle, Humphrey, and Summers (1987). In the event, a far more modest fee was charged when the Fed introduced its fee than what would have been produced by such an exercise (see Johnson, et al).

central bank overnight interest rate, adjusted for maturity and for the degree of relative collateralization. In Thailand, for example, the interest rates for borrowing under the Intraday Liquidity Facility of the Bank of Thailand are linked to the previous day's repurchase market rates.

The issue of whether the level of intraday credit matters for interest rates and inflation is a difficult one that requires some more serious research. In this context, there is some formal analysis to suggest that, in principle at least, the intraday and interday markets cannot be neatly segmented and that movements in intraday conditions and rates will indeed influence interday rates (see VanHoose, 1991).¹⁶ This would imply that, particularly as intraday markets develop, central banks may need to take this connection into account as they formulate their monetary policy. In practice, central banks generally seem, for the time being at least, of the view that any links between interday and intraday liquidity quantitatively are not yet of major importance.¹⁷ Hence, monetary policy, whatever the intermediate targets, still effectively operates in practice on end-of-day reserve balances and interday and/or overnight interest rates.

¹⁶VanHoose (1991) develops a model of bank behavior over a trading day that incorporates intraday and overnight periods. The representative bank maximizes profits and determines its reserves, earning assets, interday loans or borrowings in funds markets (24-hours or overnight) and intraday loans or borrowings. VanHoose derives the market equilibrium and shows that the equilibrium interest rates in intraday, overnight, and 24-hour funds markets are related.

¹⁷See, e.g., Ettin (1988).

B. Float

Historically, one major element of the payment system that has preoccupied monetary policy has been *float*. Float is a consequence of a time difference between the crediting of a payee's account and the debiting of a payer's account as a result of a payment transaction. The causes of float include central bank or commercial bank operational procedures (for example, crediting customers' deposit accounts when they lodge a check before the payee's bank itself receives credit for the check from the payer's bank); weaknesses in the rules or regulations (whether bank-specific or more general) governing those procedures; transportation lags in the case of paper-based payments; delayed or only partial processing of payments because of insufficient resources to finish the task by the end of the business day; and delays because of the time taken to identify and rectify processing errors.¹⁸

The central bank has a strong interest in reducing significant float, *inter alia* because of its implications for monetary management. At the operational level, the size, and even direction, of appropriate monetary operations is more difficult to set when float results in significant day-to-day volatility in the exogenous (nonmonetary policy) sources of reserves, both in the aggregate and between individual banks (see, for example, Young, 1986; and Hoel, 1975). In addition, large and variable float hinders the development of deeper and more efficient interbank and other wholesale financial markets (for instance, because of uncertainty of timing of settlement); this can slow down the shift toward use of market-based instruments of monetary policy.

¹⁸For a more detailed discussion of the nature and causes of different types of float, see Veale and Price (1994).

Forecasting and assessing demand for (and hence movements in) bank reserves, as well as broader money and credit aggregates, is further complicated by measurement issues, as well as by possible endogenous factors in float. In measurement, there are questions about the appropriate definition of “money,” in the face of significant float.¹⁹ There have also been questions about the measurement of float itself, which relate to specific accounting procedures for payments.²⁰ The issue of possible endogenous factors in float may arise in an inflationary environment, because higher inflation increases the incentives to create float and therefore may be associated with higher float—in the absence of enforceable limits on payments lags—especially in systems that are not very competitive.²¹

Hence, even leaving aside its well-known welfare effects,²² the fact that float can be very unpredictable and unstable causes serious difficulties for monetary management making policy makers anxious to reduce their average level, in relation to base money, as quickly and as much as economically feasible. Short of more fundamental reforms to speed up the payment

¹⁹For example, one suggestion in the U.S. has been that bank float needs to be subtracted from demand deposits *a second time* to measure M1 appropriately; see Liang (1986).

²⁰For example, in their work on central bank float in Russia over 1992/93, Sundararajan and Sensenbrenner (1994) note that measuring strictly payment-related float required adjusting data on incompletely processed payments for interenterprise arrears (effectively trade credit) measured through the payment system and for the difference in timing between accounting and payment of government deposits between the central bank branches and the head office.

²¹Sundararajan and Sensenbrenner (1994) found some evidence that this occurred in Russia.

²²The adverse welfare effects of float arise mainly from the waste resulting from procedures used by agents to generate and reduce float in reaction to its distributional effects.

process, several partial solutions can be (and have been) adopted to reduce float.²³ Funds *availability schedules* are an example. Under such an approach, the credit for a check (or debit for a payment order) is delayed to a time equivalent to that in which the corresponding debit (credit) would normally be processed and posted. Related is, also, *provisional crediting* of a check, where the value is credited to the payee's bank account, but the value cannot be withdrawn until the payment is finally settled. In some cases it may be possible to discourage float through imposition of appropriate pricing. When the United States adopted the Monetary Control Act of 1980, for example, measures to reduce the float by both availability schedules and pricing the remaining float were introduced (Young, 1986). The pricing involved an explicit interest charge by the Federal Reserve on the proportion of banks' reserves that could be attributed to float.

A second approach comprises attempts to *reduce some of the operational delays* and backlogs within and between banks. This can often be achieved through better staffing, improved operational procedures, and better training of processing staff, or through establishing dedicated document delivery services to reduce transportation delays where possible.

A third approach involves the affected parties offsetting the costs of float (or maximizing the gains from float) by *more active cash and short-term investment management*,

²³Account consolidation also helps to reduce float, by eliminating inter-branch float for banks.

including management of payments flows. Banks themselves may often assist their customers in such arrangements, for fear of losing them to competitors.²⁴

III. LIQUIDITY SHOCKS AND PANICS

In the absence of institutional and technological changes, routine monetary policy and operations get tested by liquidity shocks and panics (which can give rise to unexpected credit squeeze) that can be payments related. Since these can matter for inflation, it is no surprise that even those narrowly concerned only with monetary policy and operations have taken a keen interest in the measures that payment organizations and payment system regulators have, or can, put in place to reduce the likelihood that the payment system will be the origin of liquidity shocks and panics in the financial system, forcing unplanned and inflationary injection of base money by the monetary authorities. In addition, of course, the monetary authorities are concerned about preventing unwarranted deflation resulting from such shocks and panics while avoiding the moral hazard problem of being seen as a guarantor always ready to provide enough liquidity to prevent gridlock or a breakdown of the payment system or, indeed, bank failures for other reasons.

Yet, in this area, it is not clear that monetary policy can be guided only by concerns with inflation, because of the monopoly of the central bank in the production of the legal tender, the medium of exchange and the ultimate means for settling debts. Consideration of

²⁴See the discussion of cash management services, in this context, in Veale and Price (1994), pp. 156-8.

the welfare effects of liquidity shocks and panics—namely, disruption of production and inducing suboptimal intertemporal consumption patterns—may be unavoidable.

In reality, of course, it is probably the case that most liquidity shocks and panics do not originate strictly in the payment system, even though they end up putting pressure on the payment system in view of their impact on the ability of individual financial institutions to settle their obligations after clearing.²⁵ Hence measures get introduced by country authorities to improve the soundness of financial systems. Hopefully, then, as bank licensing becomes more stringent, banking supervision and governance of banks improve, and as bank branching is permitted everywhere, liquidity shocks and panics emanating from within the financial system but outside the payment system will decline in significance (frequency and severity). Liquidity shocks emanating from outside the domestic financial system—abroad or due to developments in the real sector—will continue to occur, of course.

Thus, the approach of policy makers has been straightforward and logical, namely, in addition to measures to improve the soundness of the financial system, policy makers are also instituting measures to reduce settlement risks in the payment system by ensuring payment finality²⁶ within the payment organizations, without special intervention by the central bank. Major initiatives in the industrial countries, for example, during the last two decades, can indeed be viewed in this context: *inter alia*, the steps taken by central banks to ensure that private large-value transfer systems (LVTSS)—especially net settlement systems—take

²⁵For a good introduction to the literature on banking panics, see Calomiris and Gorton (1991). In this regard, also of relevance is Freeman (1996).

²⁶Payment is final when it becomes irrevocable and unconditional.

appropriate measures to contain intraday bilateral exposures and ensure settlement finality within the systems via exposure limits (debit caps and bilateral credit limits), collateralization and loss-sharing agreements, and reduction of the time lags in settlement; the move by central banks toward real-time gross settlement (RTGS) systems; limits and collateral in central bank credit policies vis-à-vis RTGS systems; the emergence of the Lamfalussy standards as a framework for assessing and regulating private netting systems; initiative in the European Union countries to develop minimum common features for their payment systems; and cooperation among industrial countries, under the aegis of the BIS, to coordinate several aspects of their payment system policies and take other measures (such as increasing overlap of operating hours and ensuring intraday finality within systems) to reduce payment risks in foreign exchange transactions.²⁷

The above initiatives and measures contribute to reducing drastically liquidity shocks, leaving routine monetary policy and operations to worry mainly about institutional and technological changes. The central bank may increasingly have become seen as a guarantor of payments settlement because of legal obligations of the central bank, in perhaps the majority of countries, to promote monetary and financial stability, and related to this because of a growing consensus that final settlement of clearing systems should take place on the books of

²⁷For example, Japan, by extending the hours of its main foreign exchange settlement system, BOJ-NET, until 5:00 p.m. Tokyo time, has facilitated a greater overlap with CHIPS and Fedwire. The U.S. Federal Reserve has also decided to increase its hours of operating Fedwire, so that it will fully overlap with systems in Europe as well as with part of the business day in Japan. The risk in settling securities transactions has mainly been reduced by adopting systems with delivery versus payment (DVP). This is best achieved by linking book entry securities transfer and registration systems to payment systems.

the central bank;²⁸ here, again, the dominant consideration is risk—in this case systemic risk—rather than price stability as such. But, at the same time, the measures enumerated above to contain settlement risks have simultaneously reduced the chances that the central bank will be called upon to “bail out” a payment system in systemic crisis. What is good for risk is also, here, good for monetary policy.

IV. INSTITUTIONAL AND TECHNOLOGICAL CHANGES

In the context of monetary policy, institutional and technological changes basically engender changes: (1) in income velocity of money;²⁹ (2) in the money multiplier; (3) in the nature of the transmission processes and the lag structures involved, and hence the speed and predictability of the effect of central bank reserve changes on the price level; and (4) in the effectiveness of different instruments at the disposal of the monetary authorities in affecting intermediate or proximate targets of monetary policy, and hence in the optimal mix of policy instruments. In other words, institutional and technological changes in the payment system have enormous implications for monetary policy and operations. The institutional and

²⁸While banks can, as a matter of principle, settle using bilateral accounts with each other, or on the books of some private settlement agent (clearing bank), settlement by banks on the books of the central bank—whether gross or after multilateral netting—can be seen as facilitating a reduction of systemic risk. Payments using central bank money result in claims on the central bank which cannot fail (become insolvent) or have liquidity problems; from the perspective of agents other than the central bank, such payments, therefore, do not have any credit or liquidity risks associated with them.

²⁹See, e.g., Bordo, Jonung, and Siklos, (1997), for a discussion of the evidence of a long-run relationship between velocity of money and institutional and technological changes, specifically in the financial system. For a theoretical analysis of the role of technological changes in the financial sector, see, e.g., Ireland (1994).

technological changes, as we have stated already, could be endogenous, emanate from regulatory changes, or induced by incentives created by public policy changes of the authorities (monetary and nonmonetary).

A. The Long Run

It is interesting that, whereas there has been recognition that institutional and technological changes can cause changes in money demand, and ipso facto, velocity, components of the payment system get hardly mentioned explicitly in this literature. Instead the literature most frequently mentions “monetization” and “increasing financial sophistication,” both factors themselves being represented by proxies such as increasing ratio of nonbank financial assets to total financial assets, declining currency-money ratio, and rising share of labor force in non-agricultural pursuits.³⁰ Hence, we need tests that explicitly include payment system proxies for “increasing financial sophistication” and “technological changes in the financial system.”

Long-run studies support the hypothesis of a U-shaped income velocity over time, within countries, and at any time across countries, as an outcome of the opposing influences on money demand of increasing monetization and financial sophistication, and of technological

³⁰Bordo, Jonung, and Siklos (1997), for example, estimate a velocity function for five countries (Canada, U.S., U.K., Sweden, and Norway) to test the hypothesis

$$v_{it} = \beta'_0 + \beta'_1 \Phi_{it} + \beta'_2 \Omega_{it} + \epsilon_t$$

where v is velocity, i refers to the country, and t to time; Φ is a vector with elements real per capita permanent income and a proxy for the opportunity cost of holding money; and Ω is a vector with elements being the institutional factors mentioned in the text.

changes.³¹ Similarly, the money multiplier³² is expected to increase with real per capita income over time and across countries, for any given structure of real interest rates on deposits, mainly because of declining desired currency-money ratio (as the banking habit spreads and banks and bank branches multiply) and increasing financial sophistication (within banks and in the financial system as a whole) so that banks better manage (that is, better economize on) reserves. We need to isolate more clearly the effect of the payment system factors in this empirical analysis. In the meanwhile, we would posit that increasing income velocities and money multipliers are what we expect to see with most of the institutional and technological changes taking place in the payment systems around the world, although in a few cases the effect could be the opposite (such as a move from net settlement to gross real time settlement in large value transfer systems, which could cause the money multiplier to fall).

³¹See, e.g., Bordo and Jonung (1987) and Ezekiel and Adekunle (1969).

³²Recall that the money multiplier, m , is equivalent to:

$$m = \frac{\frac{C}{D} + 1}{\frac{C}{D} + \frac{R}{D}}$$

where C is currency outside banks, R is bank reserves (vault cash plus reserves held at the central bank), and D is bank deposits. Also recall that:

$$M = m \cdot H$$

where $H (=C+R)$ is high-powered or base money.

B. The Short Run

It is in short run (policy-oriented) analysis that there has been any real and explicit concern for the effect of changes in the payment system, although the theoretical and empirical literature is still relatively thin. For velocity and the money multiplier can change significantly in the short term because of major institutional and technological changes in the payment system within a short period of time. As a result, there are implications for the monetary policy decision-making process. Conceptually, four different, but closely interrelated, areas of decision making could be affected by major payment system reforms and endogenous changes, as the monetary authorities try to take into account the short-term impact on velocity and the money multiplier.³³

First are the *monetary policy target and instrument settings*—for example, the aggregate volume of reserves the central bank should supply for consistency between demand for reserves and the central bank's desired monetary policy stance; the pricing or the quantity limits in standing central bank credit facilities; and the appropriate relationship between very short-term interbank interest rates (which the central bank directly affects) and other interest rates and financial variables (over which the central bank has less direct influence). Second are the *choice of appropriate target or indicator variables* for monetary policy—for example, the

³³Recall that, abstracting from uncertainty and seasonal factors, the currency-money ratio, in the short term, will be a function of income per capita and interest rates on bank deposits. The reserve-deposit ratio, in the short term, will be a function of interest rates in money and financial markets (e.g., central bank interest rates on discounts and advances, commercial banks' interest rates on loans, and market interest rates on securities that banks keep in their portfolio); of course, in a complete analysis, all these rates are determined in a general equilibrium framework. For an illustrative partial-equilibrium empirical analysis, see Khatkhate, Galbis, and Villaneuva (1974).

relative weights (or reliability as indicators) attached to price and quantity variables (interest rates versus reserve money) while demand for the key operational quantity variable (reserve balances) is shifting. There may be effects on quantity variables at the level of the banking system, as well as at the level of the central bank's balance sheet. Third is the appropriate *design of monetary policy instruments*; for example, the design of reserve requirements or of central bank standing credit facilities, or the nature and timing of central bank market operations, might need to be adjusted. Fourth, of course, is the *monetary policy transmission mechanism* itself—for example, the efficiency with which central bank actions in respect of the supply of reserves feeds through to interest rates in different markets and subsequently through to other economic and financial variables of ultimate interest.

As an example,³⁴ in June 1987, the Swiss Interbank Clearing system (SIC) was introduced with no intraday liquidity facility but with a queuing system. When liquidity requirements were reduced in January 1988, the effects on money market rates indicated that banks had apparently introduced improved liquidity management systems, probably in response to the SIC queuing system. Thus, monetary policy turned out to be easier than expected. The Swiss National Bank also was led to modify its Lombard facility to a flexible one, *inter alia* in order to enhance its ability to respond to money market rates in a timely manner (see Rich, 1992; and Swiss National Bank, 1989).

Recent experience from various countries indicates that payment system initiatives that can be major and speedy enough to affect velocity and the money multiplier and hence influence the central bank's monetary policy decision making in the short term include:

³⁴This example is based on Lybek (1996).

(1) arrangements that reduce float; (2) development of clearinghouses and refinements of risk-reduction measures in those houses to facilitate safe and reliable netting arrangements and less frequent intraday settlement cycles; (3) moves to electronic payments (for both retail and wholesale payments); (4) centralization of commercial banks' reserve accounts at the central bank; (5) moves from net settlement to real-time gross settlement (RTGS) systems; and (6) introduction of payment instruments that reduce the use of cash or even of deposits. To reiterate, in designing their liquidity management framework and in an attempt to improve the effectiveness of monetary operations, those concerned with monetary policy and operations in central banks will find it useful to keep track of such developments in the payment system and to refine techniques to assess the short-term effects on monetary variables of policy interest.

REFERENCES

- Baliño, Tomás J. T. and Lorena M. Zamalloa, 1997, *Instruments of Monetary Management* (Washington: International Monetary Fund).
- Bank of International Settlements, 1993, *Payment Systems in the Group of Ten Countries*, prepared by the Committee on Payment and Settlement Systems of the central banks of the Group of Ten countries (Basle, December).
- Bordo, Michael D. and Lars Jonung, 1987, *The Long-Run Behavior of the Velocity of Circulation: The International Evidence* (Cambridge: Cambridge University Press).
- Bordo, Michael D., Lars Jonung, and Pierre L. Siklos, 1997, "Institutional Change and the Velocity of Money: A Century of Evidence," *Economic Inquiry*, Vol. 35 (October), pp. 710-24.
- Calomiris, Charles W. and Gary Gorton, 1991, "The Origins of Banking Panics: Models, Facts, and Bank Regulation," in *Financial Markets and Financial Crises*, edited by R. Glenn Hubbard (Chicago: University of Chicago Press), pp. 109-173.
- Ettin, Edward, 1988, "Commentary" in *Restructuring Banking and Financial Services in America*, edited by William S. Haraf and Rose Marie Kushmeider (Washington: American Enterprise Institute for Public Policy Research), pp. 113-155.
- European Monetary Institute (EMI), 1993, *Report to the Committee of Governors of the Central Banks of the Member States of the European Economic Community on Minimum Common Features for Domestic Payment Systems* (Frankfurt: Working Group on Payment Systems, November).
- _____, 1994, *Note on the EMI's Intentions with Regard to Cross-Border Payments in Stage III* (Frankfurt: Working Group on Payment Systems, November).
- _____, 1995, *Report to the Council of the European Monetary Institute on the Target System* (Frankfurt: Working Group on Payment Systems, May).
- Ezekiel, Hannan and Joseph O. Adekunle, 1969, "The Secular Behavior of Income Velocity: An International Cross-Section Study," *International Monetary Fund Staff Papers*, Vol. 16 (July), pp. 224-239.
- Freeman, Scott, 1996, "The Payments System, Liquidity, and Rediscounting," *American Economic Review*, Vol. 86 (December), pp. 1126-1138.
- Hoel, Arline, 1975, "A Primer on Federal Reserve Float, Federal Reserve Bank of New York Monthly Review Vol.57 (October), pp. 245-253.
- Ireland, Peter, 1994, "Economic Growth, Financial Evolution, and the Long-Run Behavior of Velocity," *Journal of Economic Dynamics and Control*, Vol. 18 (May/July 1994), pp. 815-48.
- Johnson, Omotunde E. G., Richard Abrams, Jean-Marc Destresse, Tonny Lybek, Nicholas M. Roberts, and Mark Swinburne, (1998), *Payment Systems, Monetary Policy, and the Role of the Central Bank* (Washington, IMF).
- Khatkhate, D. R., V. G. Galbis, and P. Villaneuva, 1974, "A Money Multiplier Model for a Developing Economy: The Venezuelan Case," *International Monetary Fund Staff Papers*, Vol. 21 (November), pp. 740-7.
- Liang, Ming-Yih, 1986, "Bank Float, Mail Float and the Definition of Money," *Journal of Banking and Finance* Vol. 10 (December), pp. 533-548.

- Lybek, Tonny, 1996, "Swiss Monetary Policy and the Introduction of SIC, 1987-89," Memorandum, Monetary and Exchange Affairs Department, IMF.
- Mengle, David, David B. Humphrey, and Bruce J. Summers, 1987, "Intraday Credit: Risk, Value, and Pricing," Federal Reserve Bank of Richmond *Economic Review* Vol. 73 (January/February), pp. 3-14.
- Rich, Georg, 1992, "Die Schweizerische Teuerung: Lehren für die Nationalbank," *Geld, Währung and Konjunktur*, Swiss National Bank, No. 10 (March), pp. 73-88.
- Sundararajan, V., and Gabriel Sensenbrenner, 1994, "Linkages Between Payment System Reforms and Monetary Policy: The Recent Experience in Russia and Other Former Soviet Union Countries," in *Frameworks for Monetary Stability*, ed. by Tomás J.T. Baliño and Carlo Cottarelli (Washington: IMF), pp. 611-54.
- Summers, Bruce J., 1994, *The Payment System: Design, Management and Supervision* (Washington: IMF, 1994).
- Swiss National Bank, 1989, "Die Geldpolitik im Jahre 1988," *Geld, Währung and Konjunktur*, Swiss National Bank, No. 4 (December), pp. 345-60.
- VanHoose, David D., 1991, "Bank Behavior, Interest Rate Determination, and Monetary Policy in a Financial System with an Intraday Federal Funds Market," *Journal of Banking and Finance* Vol. 15 (April), pp. 345-365.
- Veale, John M., and Robert W. Price, 1994, "Payment System Float and Float Management," in *The Payment System: Design, Management and Supervision* edited by Bruce J. Summers (Washington: IMF, 1994).
- Young, John E., 1986, "The Rise and Fall of Federal Reserve Float," Federal Reserve Bank of Kansas City *Economic Review* Vol. (February), pp. 28-38.