FACTORS AFFECTING MICRO, SMALL, AND MEDIUM-SIZED ENTERPRISE DEVELOPMENT IN DEVELOPING ASIA

FINDINGS FROM A PROBABILISTIC PRINCIPAL COMPONENT ANALYSIS

Shigehiro Shinozaki, Daisuke Miyakawa, and Romeo Arahan

NO. 715

February 2024

ADB ECONOMICS WORKING PAPER SERIES

ASIAN DEVELOPMENT BANK

ADB Economics Working Paper Series

Factors Affecting Micro, Small, and Medium-Sized Enterprise Development in Developing Asia: Findings from a Probabilistic Principal Component Analysis

Shigehiro Shinozaki, Daisuke Miyakawa, and Romeo Arahan

No. 715 | February 2024

Shigehiro Shinozaki (sshinozaki@adb.org) is a senior economist and Romeo Arahan (rarahan.consultant@adb.org) is a consultant at the Economic Research and Development Impact Department. Daisuke Miyakawa (damiyak@waseda.jp) is a professor at Waseda University, and chief economist of UTokyo Economic Consulting Inc. (UTEcon).

ASIAN DEVELOPMENT BANK

6 ADB Avenue, Mandaluyong City, 1550 Metro Manila, Philippines
Tel +63 2 8632 4444; Fax +63 2 8636 2444
www.adb.org

ISSN 2313-6537 (print), 2313-6545 (electronic)
Publication Stock No. WPS240032-2
DOI: http://dx.doi.org/10.22617/WPS240032-2

Note:
In this publication, "$" refers to United States dollars.

ABSTRACT

Limited data on micro, small, and medium-sized enterprises (MSMEs) make it difficult for governments to design appropriate MSME policies in Asia and the Pacific. To identify factors affecting MSME development and promote evidence-based policymaking, we propose a probabilistic principal component analysis method that works despite current data limitations. The study uses time-series MSME data collected from 25 developing member countries of the Asian Development Bank (ADB) through the Asia Small and Medium-Sized Enterprise Monitor project. The estimation results suggest that sound MSME credit markets, diversified financing options, support for new businesses and job creation, and active MSME participation in global marketplaces play a critical role in ensuring a smooth business recovery from various crises and shocks affecting developing Asia and the Pacific.

Keywords: SME development, access to finance, financial inclusion, SME policy, probabilistic principal component analysis, Southeast Asia, South Asia, Central and West Asia, the Pacific

JEL codes: D22, G20, L20, L50

1. Introduction

Developing Asian economies continue to recover from the coronavirus disease (COVID-19) pandemic that began in March 2020, although economic growth differs by country. Continuous global economic uncertainty, however, has amplified downside risks—including high inflation, currency depreciation, and global supply chain disruptions accelerated by regional political turbulence. In Southeast Asia, a recovery in tourism partly contributed to the region's 5.6% growth in 2022; but it is forecast to drop to 4.6% in 2023 given continued weak exports. In South Asia, economic and political crises in Pakistan and Sri Lanka pushed the region's growth down from 6.7% in 2022 to 5.4% in 2023. In Central and West Asia, the ongoing impact from the Russian invasion of Ukraine helped lower the region's growth from 5.1% in 2022 to a forecast 4.6% in 2023. In the Pacific, a strong post-pandemic tourism rebound energized the region's sharp economic recovery to 6.1% growth in 2022; but it is forecast to slow to 3.5% in 2023 partly due to labor shortages accelerated by emigration from small island countries to Australia and New Zealand (ADB 2023a).

Micro, small, and medium-sized enterprises (MSMEs) help drive growth across developing Asia and the Pacific, given their large share of business enterprises, job creation, and economic output. Given their impact, governments in the region have taken several policy measures to promote MSME development. They commonly promote entrepreneurial development (especially for youth and women), use of technology that encourages business innovation, expanded market access by internationalizing MSMEs, human capital and skills development, and better access to finance. But constraints on MSME development remain in most countries. These include a lack of an entrepreneurial culture, high dependence on cash transactions that stymie innovation, a large percentage of unregistered or informal businesses, limited exports or participation in global markets, skilled labor shortages, and structural problems limiting access to formal financial services for working and growth capital. This raises the question how governments can enhance policies and their implementation to promote MSME development toward more inclusive, resilient growth.

Better understanding the MSME business environment and structural problems that inhibit growth is critical before designing a feasible policy framework on MSME assistance. However, the lack of data on MSMEs makes this extremely difficult. To help governments promote evidence-based MSME policymaking, the Asian Development Bank (ADB) has, since 2020, provided benchmark indicators on MSME development and access to finance through its annual Asia Small and Medium-Sized Enterprise Monitor (ASM). As of November 2023, the ASM covers MSMEs in 25 ADB developing members in Southeast Asia, South Asia, Central and West Asia, and the Pacific. Insufficient data, however, remains a major problem.

A solid quantitative evaluation on MSME development using sufficient, accurate, and comparable data remains a challenge both nationally and regionally. Incomplete data on MSMEs led global institutions—such as the Organisation for Economic Co-operation and Development (OECD), the Economic Research Institute for ASEAN and East Asia (ERIA), and the International Trade Centre (ITC)—to propose a qualitative approach using assessment matrices for performance ratings or median comparisons based on available data to evaluate MSME development and competitiveness, both nationally and regionally.

The ASM project has also explored a new way to quantitatively identify factors affecting MSME development through its ASM database. In 2021, it developed a new trial that deals with MSME data limitations—a variant of a standard principal component analysis (PCA) that supplements some missing MSME data—a probabilistic PCA (ADB 2022). The pilot test covered 15 countries from Southeast Asia and South Asia along with a firm-level data analysis for Viet Nam. While this contributed to the new MSME development index, insufficient data limited the proposed model's ability to estimate more accurately the factors that represent MSME activities. More test-runs for additional countries are needed to produce a reliable index conducive to evidence-based policy design on MSMEs in the region. In 2023, we successfully compiled time-series MSME data covering 25 countries. With this new dataset, this study re-estimates factors that explain the MSME development path by region and country and rethinks how to develop a quantitative approach to better assess MSME development.

Section 2 summarizes the MSME landscape in developing Asia, extracted from ADB (2023b). Section 3 reviews global MSME data initiatives in Asia and the Pacific. Section 4 explains the methodology and dataset used for analysis. Section 5 discusses the estimation results in four groups—(i) all countries, (ii) Southeast Asia, (iii) South Asia, and (iv) Central and West Asia. This is followed by associated policy implications in Section 6 and conclusions in Section 7.

2. MSME Landscape in Developing Asia

MSMEs dominate the private sector in Asia and the Pacific. According to ADB (2023b), based on available data in participating countries through 2022, MSMEs in Asia and the Pacific accounted for an average 96.6% of all enterprises, 55.8% of the total workforce, and 28% of a country's economic output (Table 1). Data collected depend on the national MSME definition of each country. Most MSMEs serve small domestic markets, with many engaged in distributive trade and informal business. Cash dominates their business model and there is little incentive to grow further—categorized as "stability-oriented" firms. With a large base of informal businesses, the official MSME contribution to a country's economic output is likely well below its actual impact. Nonetheless, "growth-oriented" and innovative firms that want to expand into global markets have gradually increased across the region, although they remain a small fraction of MSMEs. Based on available data through 2022, MSME exports accounted for an average 26.3% of total export value. And MSME export growth is slowing, mainly due to the weak export environment globally. Low business diversification limits a country's growth potential, suggesting the need for creating more innovative and globalized small firms, startups, and an entrepreneurial base, both nationally and regionally.

Limited access to finance remains a chronic barrier to MSME growth. The MSME credit market remains small in Asia and the Pacific. ADB (2023b) reported that bank loans to MSMEs averaged 10.6% of a country's gross domestic product (GDP) and 22% of total bank lending. The pandemic response boosted commercial bank lending to MSMEs, provided government emergency financial assistance or strengthened new lending to MSMEs through subsidized loan programs, refinancing facilities, and special credit guarantees. Despite this, MSME nonperforming loans remained high, averaging 7.2% of total MSME bank loans in the region. The lack of alternative financing options beyond traditional bank credit limits innovation and business opportunities for viable MSMEs, startups, and entrepreneurs.

Table 1: MSMEs in Developing Asia and the Pacific
(percentage share)

	All Countries	Southeast Asia	South Asia	Central and West Asia
MSME development				
• **Number of MSMEs to total enterprises**	96.6%	98.0%	99.6%	99.2%
• **MSME employees to total employees**	55.8%	66.4%	76.6%	51.9%
• **MSME contribution to economic output**	28.0%	41.2%	17.7%	41.5%
• **MSME exports to total export value**	26.3%	13.3%	37.4%	28.3%
Access to finance (bank credit)				
• **MSME loans to national GDP**	10.6%	13.3%	5.2%	11.1%
• **MSME loans to total bank loans**	22.0%	12.3%	12.5%	33.1%
• **MSME NPLs to total MSME loans**	7.2%	5.3%	12.1%	4.3%

GDP = gross domestic product, MSME = micro, small, and medium-sized enterprise, NPL = nonperforming loan.
Notes: Reporting countries only. Data based on latest available data until 2022. Data for all countries cover 25 countries: 10 from Southeast Asia; 5 from South Asia; 7 from Central and West Asia; and 3 from the Pacific.
Source: Asia SME Monitor 2023 database.

3. Global MSME Data Initiatives

Several global initiatives are developing indices to measure specific aspects of MSMEs—such as access to markets, infrastructure, finance, skills development, use of technology and innovation, business operations and administration, competitiveness, and policy and regulatory frameworks (Table 2). Multilateral organizations such as the OECD, ERIA, ITC, and World Bank Group have been using various analytical approaches to overcome the lack of sufficient MSME data.

The OECD produces two related reports on SME development: (i) the SME and Entrepreneurship Outlook and (ii) Financing SMEs and Entrepreneurships (OECD Scoreboard). Launched in 2002, the Entrepreneurship Outlook reviews 6 dimensions with 29 subdimensions using cross sectional data for median comparison. Dimensions include (i) institutional and regulatory frameworks, (ii) market conditions, (iii) infrastructure, (iv) access to finance, (v) access to skills, and (vi) access to innovation assets (OECD 2023). The subdimensions include (i) regulations, courts and laws, land and housing, public governance, competition, and taxation; (ii) domestic markets, global markets, public procurement, and trade and investment; (iii) logistics, energy, research and development (R&D) and innovation, the internet and information and communications technology (ICT); (iv) self-funding, debt, the financial system, and alternative instruments; (v) adult literacy, the labor market, entrepreneurial culture, training, and education; and (vi) technology, R&D, organization and processes, marketing, and data. It covers OECD members, including, from Asia, Australia, Japan, New Zealand, and the Republic of Korea.

The OECD Scoreboard, launched in 2012, is an annual report focusing on trends in SME financing and policies for 48 countries. In 2022, it included 11 countries from Asia—Australia, Georgia, Indonesia, Japan, Kazakhstan, Malaysia, New Zealand, the People's Republic of China (PRC), the Republic of Korea, Thailand, and Türkiye. It reviews 5 financial dimensions with 25 subdimensions (indicators): (i) allocation and structure of bank credit to SMEs; (ii) extent of public support for SME finance; (iii) credit costs and conditions; (iv) nonbank sources of finance; and (v) financial health (OECD 2022). The OECD constructs the indicators mainly using supply-side data from standardized forms filled in by banks, other financial institutions, statistics offices, and government agencies. The core indicators include total lending (overall and SMEs), new lending (overall and SMEs), short- versus long-term SME loans, direct government SME loans, government loan guarantees, interest rates (overall and SMEs), collateral (SMEs), and bankruptcies (SMEs), among others.

The OECD and ERIA produced an ASEAN SME Policy Index in 2014 and 2018 outlining the policy landscape for SME development. It evaluates the scope and intensity of SME development policies through 8 dimensions with 25 subdimensions: (i) productivity, technology, and innovation; (ii) environmental policies targeting SMEs; (iii) access to finance; (iv) access to markets and internationalization; (v) institutional framework; (vi) legislation, regulation, and taxes; (vii) entrepreneurial education and skills; and (viii) social enterprises and inclusive entrepreneurship (OECD and ERIA 2018). These are measured in three stages: (i) planning and design; (ii) implementation; and (iii) monitoring and evaluation. Participating governments share their SME data and assess SME policies. They also conduct surveys of key stakeholders and private sector representatives to help supply missing information needed for qualitative analysis. For each subdimension, respondents score the strengths and weaknesses of current SME policies on a scale from 1 to 6, with higher scores indicating a better level of policy development and implementation.

The ITC's SME Competitiveness Outlook annually reviews SME development and financing conditions in 85 countries including several from Asia (ITC 2022). It aims to facilitate implementation of United Nations Sustainable Development Goals 8 and 9. The report produces an SME Competitiveness Index based on 3 dimensions with 39 subdimensions: (i) firm capabilities (SME's ability to manage resources under its control); (ii) business ecosystem (resources and competencies needed to enhance a firm's competitiveness); and (iii) national environment (government functionality and policy implementation). Each dimension is measured on three abilities: (i) capacity to compete (enterprise efficiency); (ii) capacity to connect (information and knowledge gathering/exploitation); and (iii) capacity to change (human and financial capital investments). The index assesses the competitive strengths and weaknesses by firm size on a 0–100 scale, analyzing time-series data obtained from secondary data including (i) the World Bank's Enterprise Surveys, Ease of Doing Business Index, and Logistics Performance Index; (ii) the International Monetary Fund's (IMF) World Economic Outlook; and (iii) the ITC's Market Access Map. Firm size classifications use the definition from World Bank Enterprise Surveys.[1] Strengths and weaknesses are measured based on a reference level of per capita GDP.

As mentioned, the World Bank Group regularly releases three related reports: (i) the International Finance Corporation (IFC) MSME Finance Gap Report; (ii) the Enterprise Surveys; and (iii) the Doing Business report. The latest IFC report was released in 2017 (with updates as needed), covering 128 countries including 29 ADB developing members (IFC 2017). Data cover general indicators such as MSME landscape, bank lending, and nonbank finance data. It estimates the potential demand for financing in emerging economies compared with current supply, and calculates the "finance gap." The report is considered a benchmark of MSME financing needs in 10 advanced economies. MSME categories include industry (manufacturing, services, or retail),

[1] The Enterprise Surveys define small firms as having 5-19 employees, medium firms 20-99, and large firms 100+ employees. https://www.enterprisesurveys.org/en/methodology.

size, and age. The World Bank Enterprise Surveys furnish the data necessary in conjunction with benchmarking for estimating the potential demand for MSME finance.

The World Bank’s Doing Business report offers thematic firm-level data for more than 130 countries covering five dimensions: (i) starting a business; (ii) hiring and firing workers; (iii) enforcing a contract; (iv) getting credit; and (v) closing a business. The report was discontinued in 2021 to be replaced by the Business Enabling Environment (BEE) Project (as of December 2022). BEE will focus on similar indicators, adding an indicator on market competition and removing the section on protecting minority investors. BEE will add analyses on digital adoption, environmental sustainability, and gender equality.

Although not focusing on MSMEs, other global initiatives on developing relevant indices include (i) the Global Competitiveness Index from the World Economic Forum (WEF); (ii) the Global Innovation Index from the World Intellectual Property Organization (WIPO); and (iii) the Global Entrepreneurship Index from the Global Entrepreneurship and Development Institute (GEDI). The Global Competitiveness Index, started in 2005, covers 141 economies including several in Asia (WEF 2019). It analyses four dimensions of competitiveness, including the enabling environment, markets, human capital, and innovation ecosystem. It uses aggregate data sourced from international organizations such as the World Bank and results of the WEF Executive Opinion Survey conducted for business executives.

The Global Innovation Index was launched in 2007 as a measuring tool for innovation in a society by using cross-sectional data for median comparisons (WIPO 2022). It is subdivided into two subindices—innovation inputs and innovation outputs. The Innovation inputs sub-index offers a snapshot of society’s enabling environment for innovation and innovative activities. Five areas are monitored: (i) institutions; (ii) human capital and research; (iii) infrastructure; (iv) market sophistication; and (v) business sophistication. The innovation outputs sub-index measures the results of innovative activities by evaluating (vi) knowledge and technology outputs and (vii) creative outputs. The average scores from both indices comprise the overall score. The 2022 edition covered 132 countries including several in Asia.

The Global Entrepreneurship Index measures a country’s entrepreneurial ecosystem, also using cross-sectional data for median comparisons (GEDI 2019). The ecosystem is the prevailing environment an entrepreneur faces. It examines entrepreneurships in terms of attitudes, abilities, and aspirations, all predicated on the society’s entrepreneurial framework—which includes market structure, infrastructure, the R&D system, financial sector, corporate sector, government, and education system. Fourteen areas are measured. A composite score is produced, and then compared nationally and regionally. The 2019 report covered 137 countries including several in Asia.

Another trial for examining factors affecting MSME development

The global MSME data initiatives reviewed above mainly use qualitative scoring methods based on national surveys with descriptive analyses and/or use median comparisons from secondary data. The limited availability of MSME data makes it difficult for direct comparisons across countries. Thus, they have tried to describe MSME conditions and identify constraints on MSME development primarily using scoring methods based on evaluation matrices that supporting governments use to design MSME policies. This study has the same purpose, but apart from these indices, applies a more quantitative approach by using panel data obtained under the Asia SME Monitor project.

Table 2: Summary of Global MSME Data Initiatives in Asia and the Pacific

Item	Asia Small and Medium-Sized Enterprise Monitor (ASM)	SME and Entrepreneurship Outlook	Financing SMEs and Entrepreneurs (OECD Scoreboard)	ASEAN SME Policy Index	SME Competitiveness Outlook	Global Competitiveness Index	Global Innovation Index	Global Entrepreneurship Index
Lead organization	**ADB**	**OECD**	**OECD**	**OECD and ERIA**	**ITC**	**WEF**	**WIPO**	**GEDI**
Year launched	2020	2002	2012	2014	2015	2005	2007	2006
Latest edition	2023	2023	2022	2018	2022	2019	2022	2019
Dimension	3	6	5	8	3	4	7	3
1	MSME development	Institutional and regulatory framework	Allocation and structure of bank credit to SMEs	Productivity, technology, and innovation	Firm capabilities	Enabling environment	Institutions	Entrepreneurial attitudes
2	Access to finance	Market conditions	Extent of public support for SME finance	Environmental policies targeting SMEs	Business ecosystem	Markets	Human capital and research	Entrepreneurial abilities
3	Policies and regulations	Infrastructure	Credit costs and conditions	Access to finance	National environment	Human capital	Infrastructure	Entrepreneurial aspirations
4		Access to finance	Nonbank sources of finance	Access to market and internationalization		Innovation ecosystem	Market sophistication	
5		Access to skills	Financial health	Institutional framework			Business sophistication	
6		Access to innovation assets		Legislation, regulation, and tax			Knowledge and technology ouputs	
7				Entrepreneurial education and skills			Creative outputs	
8				Social enterprises and inclusive entrepreneurship				
Sub-dimension	15	29	25	25	39	12	81	14
Data	Cross-sectional and time series	Cross-sectional and time series	Cross-sectional and time series	Cross-sectional (intent to create a time series)	Cross-sectional and time series	Cross-sectional and time series	Cross-sectional	Cross-sectional
Methodology	Quantitative and qualitative national surveys	Median comparison	Descriptive national surveys	Qualitative national surveys	Median comparison	Quantitative and qualitative national surveys	Median comparison	Median comparison
Participating economies	25 developing economies: (i) Southeast Asia (10); (ii) South Asia (5); (iii) Central and West Asia (7); and (iv) Pacific (3).	OECD member economies (including Australia, Japan, New Zealand, and the Republic of Korea).	48 economies (including Australia, Georgia, Indonesia, Japan, Kazakhstan, Malaysia, New Zealand, the PRC, the Republic of Korea, Thailand, and Türkiye).	10 ASEAN member states	85 economies (including Armenia, Azerbaijan, Bangladesh, Bhutan, Cambodia, Georgia, Indonesia, Kazakhstan, the Kyrgyz Republic, the Lao PDR, Mongolia, Myanmar, Nepal, Pakistan, the Philippines, Tajikistan, Timor-Leste, Türkiye, and Viet Nam).	141 economies (including Armenia; Australia; Azerbaijan; Bangladesh; Brunei Darussalam; Cambodia; the PRC; Georgia; Hong Kong, China; Indonesia; India; Japan; Kazakhstan; the Republic of Korea; the Kyrgyz Republic; the Lao PDR; Malaysia; Mongolia; Nepal; Pakistan; the Philippines; Singapore; Sri Lanka; Tajikistan; Thailand; Türkiye; and Viet Nam).	132 economies (including Armenia; Australia; Azerbaijan; Bangladesh; Brunei Darussalam; Cambodia; the PRC; Georgia; Hong Kong, China; Indonesia; India; Japan; Kazakhstan; the Republic of Korea; the Kyrgyz Republic; the Lao PDR; Malaysia; Mongolia; Nepal; Pakistan; the Philippines; Singapore; Sri Lanka; Tajikistan; Thailand; Türkiye; Uzbekistan; and Viet Nam).	137 economies (including Armenia; Australia; Azerbaijan; Bangladesh; Brunei Darussalam; Cambodia; the PRC; Georgia; Hong Kong, China; Indonesia; India; Japan; Kazakhstan; the Republic of Korea; the Kyrgyz Republic; the Lao PDR; Malaysia; Myanmar; Pakistan; the Philippines; Singapore; Tajikistan; Thailand; Türkiye; and Viet Nam).
Remarks	Country and regional reviews	SME performance and the degree of entrepreneurship	Finance and entrepreneurship scoreboard	Policy landscape relates to SME development and policy implementation	SME competitiveness at the macro level	Competitiveness and economic productivity	Innovation trends and analysis	Entrepreneurship ecosystem

ADB = Asian Development Bank; ASEAN = Association of Southeast Asian Nations; ERIA = Economic Research Institute for ASEAN and East Asia; GEDI = Global Entrepreneurship and Development Institute; ITC = International Trade Centre; Lao PDR = Lao People's Democratic Republic; MSME = micro, small, and medium-sized enterprise; OECD = Organisation for Economic Co-operation and Development; PRC = People's Republic of China; WEF = World Economic Forum; WIPO = World Intellectual Property Organization.
Source: Authors. Updates as of 30 October 2023.

4. Methodology and Dataset

Given MSME data limitations, a model was developed using a probabilistic principal component analysis (P-PCA), combined with estimating the expectation-maximization (EM) algorithm to compensate for missing data. The exercise used the ADB Asia SME Monitor 2023 database.[2] The following sections explain the structure of the data used and the model specifications.

4.1. Data Structure

The Asia SME Monitor database stores various MSME-related aggregate variables, covering 25 countries as of November 2023—10 in Southeast Asia (Brunei Darussalam, Cambodia, Indonesia, the Lao People's Democratic Republic [Lao PDR], Malaysia, Myanmar, the Philippines, Singapore, Thailand, and Viet Nam); 5 in South Asia (Bangladesh, India, Nepal, Pakistan, and Sri Lanka); 7 in Central and West Asia (Armenia, Azerbaijan, Georgia, Kazakhstan, the Kyrgyz Republic, Tajikistan, and Uzbekistan), and 3 in the Pacific (Fiji, Papua New Guinea, and Samoa).[3]

The database covers three dimensions: (i) the MSME landscape—14 data categories including the number of MSMEs, those employed by MSMEs, contribution to economic output (whether in GDP or gross value added), and MSME export/import values; (ii) MSME access to bank credit—8 data categories including MSME bank loans outstanding, nonperforming MSME loans, and guaranteed loans; and (iii) MSME access to nonbank and market-based finance—8 data categories including nonbank finance institution (NBFI) finance (including microfinance institutions, finance companies, credit cooperatives, and pawnshops), nonperforming financing, and the market capitalization MSMEs can tap. All local currency data were converted into US dollars, referring to end-of-year currency rates from the IMF International Financial Statistics for designated years. The data covers 2007–2022.

Data with sufficient observations are used as independent variables to estimate the latent variable "MSME development" under the P-PCA model.[4] There are two groups of variables incorporated into the model: nonfinance and finance data. For nonfinance data, the variable "number of MSMEs" is the number of enterprises meeting the MSME criteria for each country and year. It indicates the net data provided by national statistics agencies and does not show details of a firm's "scrap-and-build" conditions, but an increased number roughly suggests newly created businesses. The variable "number of employees" measures the number of workers employed by the MSMEs in each country and year. "MSME output" measures the sum of value-added produced by MSMEs in each country and year. The variables "MSME exports" and "MSME imports" show the value of products exported and imported by MSMEs in each country and year.

Finance-related variables measure the state of corporate financing in each country and year. The variables "MSME loans outstanding" and "nonperforming MSME loans" correspond to the outstanding amounts of bank loans to MSMEs and the amount of nonperforming loans. Given that data on MSME bank loans are unavailable for some countries, "bank loans outstanding" and "nonperforming bank loans" are included in the datasets, as these include MSME borrowers. The variables "NBFI loans outstanding" and "nonperforming NBFI loans" also refer to loans from nonbank finance institutions available for MSMEs. The variable "market capitalization" is the

---

[2] ADB Asia SME Monitor 2023 database. https://data.adb.org/dataset/2023-asia-small-and-medium-sized-enterprise-monitor.

[3] Myanmar was excluded after 2020 for the data update. Effective 1 February 2021, ADB placed a temporary hold on sovereign project disbursements and new contracts in Myanmar.

[4] In the case of an extremely small number of observations, P-PCA estimates display "errors."

market value of listed companies on dedicated MSME market boards or equity markets that MSMEs can tap in each country and year. For countries without dedicated MSME markets or where MSME market data are unavailable, main market data are used, given that equity financing is an important external financing source for MSMEs.

Appendix 1 summarizes each variable for each country. The number of available variables varies by country. Although the range of available data has largely improved compared with the 2021 pilot test, even country-level aggregate variables are not commonly available for all countries. In addition, the number of observations for each variable varies for each country, creating missing values in certain variables (Appendix 2). These facts support the use of the P-PCA method for this study.

4.2. Regression Models

Probabilistic principal component analysis

Missing data is always an issue for data analysis. The P-PCA, while a derivation of principal component analysis, has been used to solve problems and issues relating to missing data across different sectors in social science and engineering by analyzing, predicting, or detecting variables of interest.

Several studies have focused on the effectiveness of P-PCA in imputing for missing data by running experiments and comparing results by using other methods of imputing missing data. For instance, Hegde et al. (2019) conducted an experiment where some data points were deliberately omitted. P-PCA was used to estimate whether the predicted values would be closest to the original data and then compared the results using multiple imputation using chained equations (MICE). The experiment showed that P-PCA was the better statistical tool for imputing missing completely at random (MCAR) data than MICE. Another experiment run by Jenelius and Koutsopoulos (2017) tried to predict taxi travel times in urban networks using P-PCA and k-nearest neighbors. It revealed that the results of P-PCA provided more accurate predictions.

P-PCA was also used to input missing data for image analysis and reconstruction. Yu et al. (2010) utilized P-PCA to impute missing data that would help reconstruct 3D images. Employing an algorithm using P-PCA and expectation maximization proved an effective way to reconstruct 3D images. Cao, Liu, and Yang (2008) used P-PCA to help detect small infrared targets by helping map the input vector from the image onto a subspace. It better predicted the possibility of the input being a target.

Other experiments utilized P-PCA to detect and filter out abnormal data and outliers. This was pivotal in addressing data issues, such as identifiability issues and removal of bias in the analysis. Qu et al. (2009) used robust PCA to filter out abnormal traffic flow data and compared it to other methods such as the nearest/mean historical imputation method and local interpolation/ regression method. Compared with other methods, it showed P-PCA reduced the root-mean-square imputation error by at least 25%. Chen, Martin, and Montague (2009) successfully used P-PCA as a tool to detect outliers and were able to run contribution analysis after yielding the missing data. Their research showed that P-PCA was critical in identifying the source of the outliers, thereby improving their analysis. Xiang, Zhong, and Gao (2015) used P-PCA to better detect rolling element bearing faults and then conducted spectral kurtosis to determine the optimal center frequency and bandwidth. Ma et al. (2021) used P-PCA to create a base model to detect anomalies and identify structural damage in buildings. The experiment proved that P-PCA is effective in recovering missing data to conduct the analysis.

Advances in data science and analysis require using machine learning and neural networks to support sophisticated processes. This allows P-PCA to be used in conjunction with more advanced tools to achieve research goals. Dixit, Bhagat, and Dangi (2022) were successful in integrating P-PCA in detecting fake news. P-PCA was utilized to improve the filtering process of news after initial manual filtering. By automating the process, the authors were able to detect and classify fake news using long short-term memory networks.

Overall, P-PCA has been an effective tool in imputing missing data that allows for predictive modelling, anomaly and outlier detection, and in improving data analysis across different sectors in social sciences and engineering. It has worked better than other missing data estimation techniques. Combined with other sophisticated data analysis methodologies, P-PCA helps increase accuracy and reduces errors.

Concept of P-PCA

This section provides a detailed explanation of the P-PCA model developed for this study, referring to Tipping and Bishop (1999), Bishop (2006), and Hastie et al. (2009).[5] The P-PCA is a variant of a standard PCA that allows for some missing data, assuming that every observed data $x \in \mathbb{R}^d$ correspond to a latent variable $z \in \mathbb{R}^{d'}$ and are generated by the following linear model:

$$x = Wz + \mu + \epsilon, \tag{1}$$

where matrix W$\in \mathbb{R}^{d\times d'}$ relates the latent variable to the observed data, μ ($\in \mathbb{R}^d$) is the mean of this model, and ε is the noise. The distribution of z is the $k-$dimensional standard Gaussian $N(0, I)$, while ϵ comes from the Gaussian $N(0, \sigma^2 I)$. When there is n observed data $\{x_i\}_{i=1}^n$, the latent variable and a noise corresponding to x_i are written as z_i and ϵ_i, respectively. For simplicity, $X = (x_1, x_2 \cdots, x_n)^T \in \mathbb{R}^{n\times d}$ and $Z = (z_1, z_2 \cdots, z_n)^T \in \mathbb{R}^{n\times d'}$, and each x_i is regarded as identically and independently sampled from model (1). Thus, model (1) assumes the observed data is realized by the low dimensional $(d' < d)$ latent variable. The goal is to find optimal parameters (μ, W, σ^2) to maximize the posterior likelihood. Before applying the analysis, X is regularized into mean 0 and variance 1 for each column.

Under these premises, the observed variable s_i follows its marginal distribution $x_i \sim N(\mu, WW^T + \sigma^2 I)$ (independent and identically). Thus, the following log likelihood function is generated, where X is regularized as the zero-mean in the following transformation:

$$\begin{aligned}\mathcal{L} &= \sum_{i=1}^{n} -\frac{1}{2}\left[\log\det\left(WW^\top + \sigma^2 I\right) + (x_i - \mu)^\top \left(WW^\top + \sigma^2 I\right)^{-1}(x_i - \mu)\right] \\ &= -\frac{1}{2}\left[n\log\det\left(WW^\top + \sigma^2 I\right) + n\mu^\top \left(WW^\top + \sigma^2 I\right)^{-1}\mu + \sum_{i=1}^{n} x_i^\top \left(WW^\top + \sigma^2 I\right)^{-1} x_i\right],\end{aligned} \tag{2}$$

where the independent terms for the maximum likelihood estimation are omitted.

According to Tipping and Bishop (1999), the optimal parameters that attain the maximal of L can be explicitly written. Here, the eigenvalue decomposition of the covariance matrix of X is used.

[5] The methodology is explained in ADB (2022, pp. 4–7). As the same analytical estimation process was followed using expanded datasets through 2022, the same explanation applies here.

Let $(v_1, \lambda 1), \cdots, (v_d, \lambda_d)$ be sets of eigenvector and eigenvalue of $X^T X$ sorted in order of increasing eigenvalues. With $U = (v_1, \cdots, v_d)$ and $\Lambda = (\lambda_1, \cdots, \lambda_d)$, these notations bring the following solution for the maximum likelihood estimation:

$$\begin{cases} \mu_* = 0 \\ W_* = U(\Lambda - \sigma^2 I)^{\frac{1}{2}} R \quad \left(R \ (\in \mathbb{R}^{d' \times d'}) \text{ is an arbitrary rotation matrix}\right) \\ \sigma_*^2 = \dfrac{1}{d-d'} \displaystyle\sum_{i=d'+1}^{d} \lambda_i. \end{cases} \tag{3}$$

Using expectation–maximization algorithm

Apart from the explicit solution to equation (3), there are several useful iterative algorithms to solve optimization problems. The gradient descent method is probably the most popular algorithm for optimization. The expectation-maximization (EM) algorithm introduced here assures that L does not decrease in each step.

In the following exposition, we denote a set of parameters (μ, W, σ^2) as θ, and subscript the parameters of the k-th iteration as $\theta_k (k = 1,2, \cdots)$. However, a more generalized setting is considered where the random variables *x* and *z* follow a joint distribution $p(x, z|\Theta)$, but only x can be observed. We draw n data $\{x_i\}_{i=1}^{n}$ identically and independently from $p(x, z|\theta)$. To apply the maximum likelihood estimation to derive θ_k, the objective function can be written as follows:

$$\begin{aligned} \mathcal{L}(\theta) &= \log p(X|\theta) \\ &= \int p(Z|X, \theta_k) \log p(X|\theta) \ \mathrm{d}Z \\ &= \int p(Z|X, \theta_k) \log \frac{p(X, Z|\theta)}{p(Z|X, \theta_k)} \ \mathrm{d}Z + \int \log p(Z|X, \theta_k) \ \log \frac{p(Z|X, \theta_k)}{p(Z|X, \theta)} \ \mathrm{d}Z \\ &= \int p(Z|X, \theta_k) \log \frac{p(X, Z|\theta)}{p(Z|X, \theta_k)} \ \mathrm{d}Z + \mathrm{KL}\left(p(Z|X, \theta_k) \parallel p(Z|X, \theta)\right). \end{aligned}$$

The EM algorithm aims to maximize the first term so that the following equation holds:

$$\theta_{k+1} = \arg\min_{\theta} \int p(Z|X, \theta_k) \log \frac{p(X, Z|\theta)}{p(Z|X, \theta_k)} \ \mathrm{d}Z = \arg\min_{\theta} \int p(Z|X, \theta_k) \log p(X, Z|\theta) \ \mathrm{d}Z$$

The second term is no less than zero and attains its minimum at $\theta = \theta_k$. Thus, the EM algorithm assures that $L(\theta_{k+1}) \geq L(\theta_k)$ holds for all $k = 1,2, \cdots$ as follows:

$$Q_k(\theta) = \int p(Z|X, \theta_k) \log p(X, Z|\theta) \ \mathrm{d}Z = \sum_{i=1}^{n} \int p(Z|X, \theta_k) \log p(x_i, z_i|\theta) \ \mathrm{d}Z = \sum_{i=1}^{n} \int p(z|x_i, \theta_k) \log p(x_i, z_i|\theta) \ \mathrm{d}z.$$

The second equality holds as (x_i, z_i) $(i = 1, \cdots, n)$ is independent.

In summary, the EM algorithm alternately repeats two steps: an expectation step for log $p(X, Z|\theta)$ with regard to $p(Z|X, \theta_k)$ and to calculate the *k*-th target function $Q_k(\theta)$, and a maximization step to maximize it. Although there is no guarantee of obtaining a global optimal solution, convergence of its likelihood is guaranteed by its derivation. This method is particularly useful for estimating parameters in latent variable models where the optimization of simultaneous distributions is difficult.

Applying the EM algorithm to the probabilistic PCA model, the simultaneous distributions can be written as follows:

$$p(x, z|\mu, W, \sigma^2) = \frac{1}{\sqrt{(2\pi\sigma^2)^d}} \exp\left(-\frac{\|x - Wz - \mu\|_2^2}{2\sigma^2}\right) \cdot \frac{1}{\sqrt{(2\pi)^k}} \exp\left(-\frac{\|z\|_2^2}{2}\right).$$

Then, the conditional distribution $p(z|x_i, \theta_k)$ of z with *k*−th parameters is given by the following:

$$p(z|x_i, \theta_k) \propto \exp\left(-\frac{1}{2\sigma_k^2}\left(z^\top (W_k^\top W_k + \sigma^2 I)z - 2(x_i - \mu_k)^\top W_k z\right)\right)$$
$$\therefore z|x_i, \theta_k \sim \mathcal{N}\left((W_k^\top W_k + \sigma_k^2 I)^{-1} W_k^\top (x_i - \mu_k), \sigma_k^2 (W_k^\top W_k + \sigma_k^2 I)^{-1}\right)$$

This object means that the mean $< z_i >$ and the covariance $< z_i z_i^T >$ of z under $p(z|x_i, \theta_k)$ can be written respectively as follows, with $\left(W_k^T W_k + \sigma^2 I\right)^{-1}$ denoted as M_k:

$$<z_i> = M_k W_k^\top (x_i - \mu_k), \quad <z_i z_i^\top> = \sigma_k^2 M_k + <z_i><z_i>^\top$$

By extracting terms which relate to θ from Q_k, we get the *k*−th target function:

$$\begin{aligned}
&\sum_{i=1}^{n} \int p(z|x_i, \theta_k) \log p(x_i, z_i|\theta) \, \mathrm{d}z - (\text{irrelevant terms}) \\
&= \frac{1}{2} \sum_{i=1}^{n} \int p(z|x_i, \theta_k) \left[\log \sigma^2 + z^\top z + \frac{1}{\sigma^2}\left\{(x_i - \mu)^\top (x_i - \mu) - 2(x_i - \mu)^\top W z + z_i^\top W^\top W z_i\right\}\right] \, \mathrm{d}z \\
&= \frac{1}{2}\left[\log \sigma^2 + \mathrm{tr}(<z_i z_i^\top>) + \frac{1}{\sigma^2}\left\{(x_i - \mu)^\top (x_i - \mu) - 2(x_i - \mu)^\top W <z_i> + \mathrm{tr}(W^\top W <z_i z_i^\top>)\right\}\right].
\end{aligned}$$

Finally, θ_{k+1} is calculated by differentiating the target function by (μ, W, σ^2) and finding a unique stationary point:

$$\begin{cases} \mu_{k+1} = 0 \\ W_{k+1} = \left(\sum_{i=1}^{n} x_i{<}z_i{>}^\top \right) \cdot \left(\sum_{i=1}^{n} {<}z_i z_i^\top{>} \right)^{-1} \\ \sigma_{k+1}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(\|x_i\|_2^2 - 2x_i^\top W_{k+1}{<}z_i{>} + \mathrm{tr}(W_{k+1}^\top W_{k+1}{<}z_i z_i^\top{>}) \right) \end{cases},$$

under the condition $\mu_0 = 0$.

When the set of data X is missing some values, each x_i is decomposed into the following two terms for easier explanation:

$$x_i = I_i^s s_i + I_i^t t_i.$$

The two variables s_i and t_i correspond to the observed coordinates and missing coordinates, respectively. When x_i consists of u observed coordinates and v missing coordinates, $(s_i)_j$ is the $j-$th observed coordinates of x_i, and $(t_i)_j$ is the $j-$th missing coordinates of x_i. Therefore, s_i and t_i are u-dimensional and v-dimensional, while $I_i^s (\in \mathbb{R}^{d\times u})$ and $I_i^t (\in \mathbb{R}^{d\times v})$ are defined as follows:

$$(I_i^s)_{lj} = \begin{cases} 1 & (\text{if } j-\text{th observed coordinate of } x_i \text{ is } x_l) \\ 0 & (\text{otherwise}), \end{cases}$$
$$(I_i^t)_{lj} = \begin{cases} 1 & (\text{if } j-\text{th missing coordinate of } x_i \text{ is } x_l) \\ 0 & (\text{otherwise}). \end{cases}$$

Then, the simultaneous distribution of s, t, z under fixed I^s and I^t can be written as follows:

$$p(s,t,z|\mu,W,\sigma^2) = \frac{1}{\sqrt{(2\pi\sigma^2)^d}} \exp\left(-\frac{\|I^s s + I^t t - Wz - \mu\|_2^2}{2\sigma^2} \right) \cdot \frac{1}{\sqrt{(2\pi)^k}} \exp\left(-\frac{\|z\|_2^2}{2} \right).$$

Also, there are conditional distributions about z and t under the observed s:

$$p(t,z|s,\mu,W,\sigma^2) \propto \exp\left(-\frac{1}{2\sigma^2} \begin{bmatrix} z^\top t^\top \end{bmatrix} \underbrace{\begin{bmatrix} W^\top W + \sigma^2 I & -W^\top I^t \\ -(I^t)^\top & I \end{bmatrix}}_{=:D} \begin{bmatrix} z \\ t \end{bmatrix} + \frac{1}{\sigma^2} \underbrace{\left(\begin{bmatrix} W^\top \\ -(I^t)^\top \end{bmatrix} (I^s s - \mu) \right)}_{=:m}^\top \begin{bmatrix} z \\ t \end{bmatrix} \right).$$

We define and calculate m and D with the mean and variance of t and z under the fixed s. Using these distribution functions, we can derive the EM algorithm for data with missing values by regarding both t and z as latent variables.

5. Estimation Results

This section presents the estimation results for (i) all 25 countries, (ii) Southeast Asia, (iii) South Asia; and (iv) Central and West Asia. A regional estimation for the Pacific was not conducted as there were only 3 countries included in the model. A robustness test was done by applying another variant of P-PCA to aggregate data of the 25 countries (Appendix 4).

5.1. All Countries

The P-PCA was applied to country-level panel data of 25 countries to see the time-series dynamics of MSME development in Asia and the Pacific. Three factors were obtained—principal component (PC)1 to PC3 (Figure 1, Table 3). PC1 makes the largest contribution to the variation in country-level panel data (59%), followed by PC2 (15%) and PC3 (7%), explaining 80% in total (Table 4). Factor loadings are sorted in descending order (Table A3.1). A darker red color indicates a positive impact to the trend in the principal component, while a darker blue means a negative impact. Each factor is orthogonal to each other and related to each variable with specific factor loadings. Key factors that form each PC can thus be extracted.

PC1-PC3 formed three different trend curves on MSME development (Figure 1). PC1 traces a low line until the middle of the sample period, then rises from 2015 until it slows after 2020. It suggests that MSMEs felt the effects of the aftermath of the 2008–2009 global financial crisis (GFC) until 2014, when recovery accelerated until development slowed after the COVID-19 pandemic spread in 2020. Overall, it indicates “a slow recovery against the shocks.” PC2 remains low until 2011, moves up until the 2015 peak, then declines afterward. It suggests that MSMEs made relatively rapid recovery from the GFC, then decelerated development around the latter part of the 2014–2016 Russian Financial Crisis (RFC), and shifted to the negative after 2019, accelerated by the 2020–2021 COVID-19 pandemic. It indicates “relatively fast recovery against the shocks but sensitive to the shocks.” The PC3 curve is more complicated, rising soon after the GFC with its first peak in 2011, bottoming out in 2016 (RFC), and rising again afterward. It suggests “a quick recovery against the shocks but very sensitive to the shocks.”

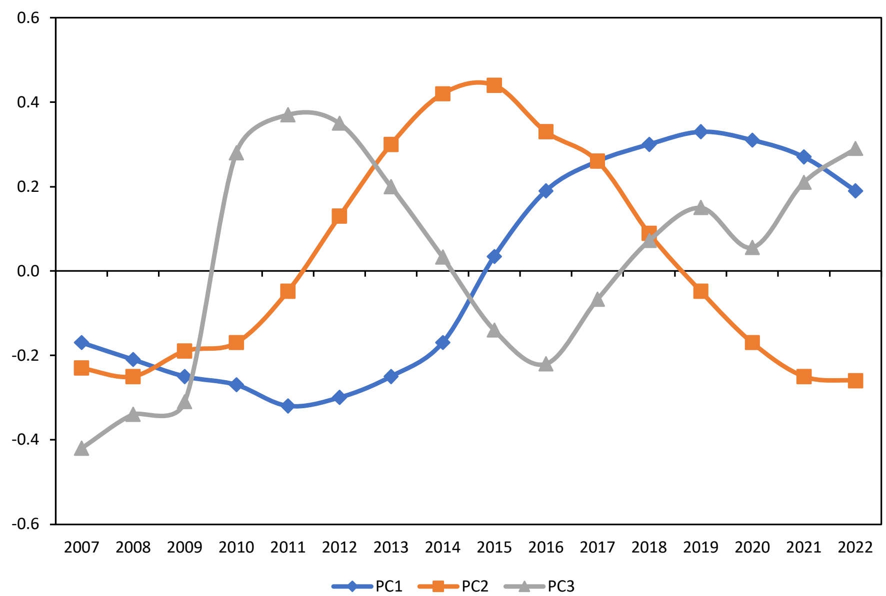
In PC1 (**slow recovery**), the main factors slowing MSME development were nonperforming loans by banks, NBFIs, and for MSMEs (e.g., Pakistan, Kazakhstan, Brunei Darussalam, Malaysia, and Viet Nam). Negative factor loadings also indicated “MSME loans" (e.g., Tajikistan, Kazakhstan, Papua New Guinea, and Georgia), which means that increased MSME loans in some countries lowered the level of MSME development under PC1. This suggests that the delivery of low quality MSME loans with increased nonperforming loans in the countries mentioned likely contributed to MSMEs’ slow GFC recovery until 2014. In contrast, the main factors that boosted MSME development were (i) bank loans (e.g., the Philippines, Fiji, the Kyrgyz Republic, and India), (ii) number of MSMEs (e.g., Indonesia, Nepal, Georgia, Viet Nam, and the Kyrgyz Republic), and (iii) MSME output (e.g., Indonesia and Pakistan). This suggests that steady delivery of bank loans to businesses likely catalyzed the increase in number of MSMEs (new small business creation) and output, bringing MSMEs back to their growth path after 2015.

In PC2 (**relatively fast recovery**), key drivers that lowered MSME development were also nonperforming loans by banks, NBFIs, and for MSMEs (e.g., Georgia, Fiji, Bangladesh, Uzbekistan, the Philippines, Thailand, Papua New Guinea, the Kyrgyz Republic, Viet Nam, Cambodia, and Sri Lanka). On the other hand, factors that accelerated MSME development were (i) number of MSME employees (e.g., India, Uzbekistan, Viet Nam, Georgia, Tajikistan, Malaysia, Indonesia, and the Philippines), (ii) MSME output (e.g., Tajikistan, Uzbekistan, and Georgia), (iii) MSME loans (e.g., the Lao PDR, the Philippines, Thailand, and Malaysia), and (iv) equity market capitalization (e.g., the Lao PDR, Sri Lanka, and Pakistan). PC2 suggests that nonperforming MSME loans likely slowed MSME development after the GFC and the COVID-19 pandemic, while improved delivery of MSME loans and the recovery of market-based finance resulted in a better environment for new MSME jobs and enhanced output. This likely supported the relatively fast recovery and growth of MSME businesses after the GFC.

In PC3 (**quick but sensitive recovery**), nonperforming loans by banks and for MSMEs (e.g., the Lao PDR, Indonesia, and Tajikistan) were again the main drivers slowing MSME development. Factors that raised the MSME development level were (i) bank loans (e.g., Malaysia, Thailand, Armenia, and India), (ii) market capitalization (e.g., Indonesia, Papua New Guinea, and Bangladesh), (iii) MSME output (e.g., Malaysia, Kazakhstan, Azerbaijan, Georgia, and the Kyrgyz Republic), and (iv) MSME exports and/or imports (e.g., Indonesia and the Kyrgyz Republic). Although PC3 showed a small contribution to explaining MSME development, it suggests that expanded bank lending and capital markets likely helped the rapid recovery of MSME exports and output and quickly accelerated MSME development. But due to weak financial markets and international trade for MSMEs, it remained highly sensitive to shocks like the GFC, RFC, and COVID-19 pandemic.

Overall, the estimation results for all countries show that sound MSME credit markets, diversified financing options (market-based finance), support for new business development and job creation, along with active MSME participation in global markets play a critical role for the smooth recovery from crises and shocks in developing Asia and the Pacific. Sound, resilient finance sector development is indispensable for sustainable MSME growth nationally.

Figure 1: Time Series Plots of Estimated Principal Components—All Countries



PC = principal component.
Source: Calculated based on ADB Asia SME Monitor 2023 database.

Table 3: Time Series Plots of Estimated Principal Components—All Countries

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
PC1	-0.17	-0.21	-0.25	-0.27	-0.32	-0.30	-0.25	-0.17	0.03	0.19	0.26	0.30	0.33	0.31	0.27	0.19
PC2	-0.23	-0.25	-0.19	-0.17	-0.05	0.13	0.30	0.42	0.44	0.33	0.26	0.09	-0.05	-0.17	-0.25	-0.26
PC3	-0.42	-0.34	-0.31	0.28	0.37	0.35	0.20	0.03	-0.14	-0.22	-0.07	0.07	0.15	0.06	0.21	0.29

PC = principal component.
Source: Calculated based on ADB Asia SME Monitor 2023 database.

Table 4: Contribution of Each Estimated Principal Component—All Countries

Item	PC1	PC2	PC3
Contribution ratio	0.59	0.15	0.07
Cumulative contribution rate	0.59	0.74	0.80

PC = principal component.
Source: Calculated based on ADB Asia SME Monitor 2023 database.

5.2. Southeast Asia

For Southeast Asia, three PC factors (PC1-PC3) were also obtained (Figure 2, Table 5). PC1 makes the largest contribution to the variation in country data (59%), followed by PC2 (16%) and PC3 (7%), explaining 82% in total (Table 6). Factor loadings are sorted in descending order (Table A3.2).

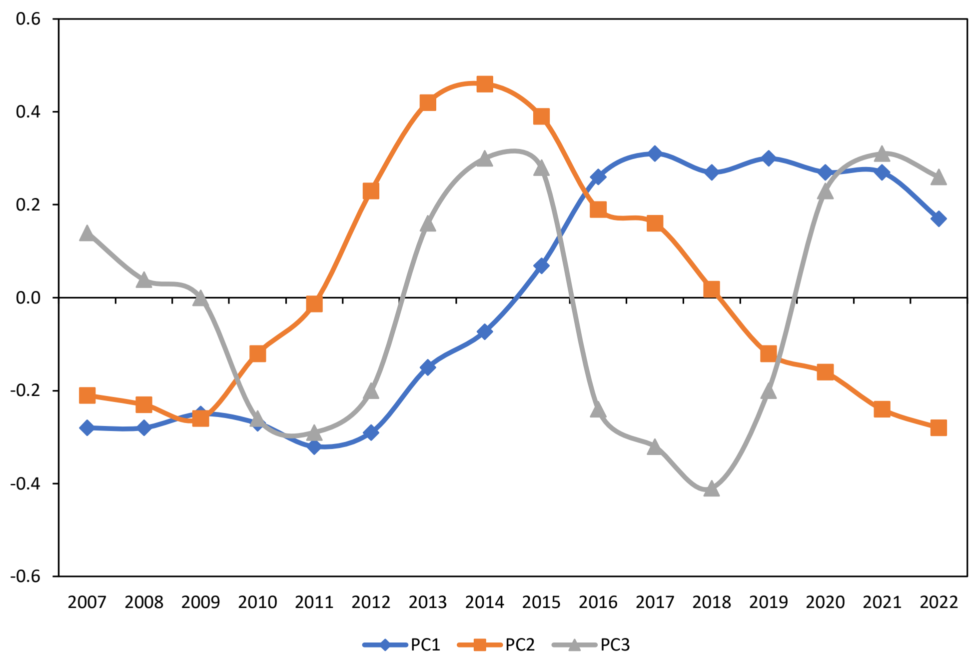
PC1-PC3 in Southeast Asia followed similar trend curves as in “all countries” but with somewhat more complicated shapes (Figure 2). PC1 remains low until 2014, then rises to a 2017 peak, before decelerating growth with a drop in 2022. It indicates a slow recovery path from the GFC. PC2 also remains low until 2011, moves up until the peak in 2014, and then declines afterward—reaching its peak a year earlier than in “all countries.” It indicates a fast recovery from the GFC but sensitive to shocks such as the RFC and the pandemic. PC3 generated a complicated shape with frequent ups and downs during 2007–2022. It bottoms out in 2011, moves up to a peak in 2014, then falls until 2018 before rising again to a peak in 2021. It indicates that MSME development is very sensitive to shocks such as the GFC, RFC, and the pandemic, while also quickly recovering.

In PC1 (**slow recovery**), the negative curve until 2014 is explained by negative factor loadings denoted by nonperforming loans by banks and for MSMEs (e.g., Brunei Darussalam, Malaysia, Viet Nam, and the Philippines). In contrast, the positive curve after 2015 is explained by positive factor loadings denoted by (i) bank loans and NBFI loans (e.g., the Philippines, Viet Nam, the Lao PDR, Thailand, and Singapore), (ii) number of MSMEs (e.g., Indonesia and Viet Nam), and (iii) MSME output (e.g., Indonesia, Brunei Darussalam, and Thailand). PC1 suggests that increased nonperforming loans likely contributed to their slow recovery from the GFC. But improved lending by banks and NBFIs likely facilitated new small business creation and a rebound in output, allowing a return to development growth after 2015.

In PC2 (**fast recovery**), the negative curve until 2011 and after 2019 is also explained by negative factor loadings denoted by nonperforming loans by banks, NBFIs, and for MSMEs (e.g., Thailand, Viet Nam, Myanmar, Cambodia, and the Philippines). The positive curve peaking in 2014 is explained by positive factor loadings denoted by (i) MSME loans (e.g., the Lao PDR, the Philippines, Thailand, and Malaysia), (ii) market capitalization (e.g., the Lao PDR, Singapore, and Thailand), and (iii) number of MSME employees (e.g., Viet Nam, Malaysia, the Philippines, and Indonesia). PC2 suggests that the high level of nonperforming loans in the finance sector likely impeded MSME development around the GFC and the pandemic, while expanded MSME lending, the capital market recovery (including dedicated MSME equity markets such as *Catalist* in Singapore and *mai* in Thailand), along with more MSME jobs likely helped the relatively fast MSME development post GFC.

In PC3 (**sensitive recovery**), the negative curve around two points in 2011 and 2018 is explained by negative factor loadings denoted by (i) nonperforming loans by banks, NBFIs, and for MSMEs (e.g., the Lao PDR, Singapore, Brunei Darussalam, Thailand, and Viet Nam) and (ii) MSME loans (e.g., Thailand, the Philippines, and the Lao PDR). The positive curve around two points in 2014 and 2021 is explained by positive factor loadings denoted by (i) loans by banks, NBFIs, and MSME lending (e.g., Singapore, Brunei Darussalam, and Cambodia) and (ii) market capitalization (e.g., Malaysia [ACE and LEAP markets] and the Philippines [SME Board]). PC3 generated a different curve than for "all countries", more pronounced in the effect of access to finance. It suggests that low quality MSME loans with increased nonperforming loans in countries such as Thailand and the Lao PDR likely kept MSME development suppressed in the region (especially in 2016–2018 amid the global economic slowdown). But diversified financing options from bank credit along with nonbank and market-based finance likely helped MSMEs recover from the shocks smoothly, while volatile financial markets held back resilience.

Figure 2: Time Series Plots of Estimated Principal Components—Southeast Asia



PC = principal component.
Source: Calculated based on ADB Asia SME Monitor 2023 database.

Table 5: Time Series Plots of Estimated Principal Components—Southeast Asia

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
PC1	-0.28	-0.28	-0.25	-0.27	-0.32	-0.29	-0.15	-0.07	0.07	0.26	0.31	0.27	0.30	0.27	0.27	0.17
PC2	-0.21	-0.23	-0.26	-0.12	-0.01	0.23	0.42	0.46	0.39	0.19	0.16	0.02	-0.12	-0.16	-0.24	-0.28
PC3	0.14	0.04	0.00	-0.26	-0.29	-0.20	0.16	0.30	0.28	-0.24	-0.32	-0.41	-0.20	0.23	0.31	0.26

PC = principal component.
Source: Calculated based on ADB Asia SME Monitor 2023 database.

Table 6: Contribution of Each Estimated Principal Component—Southeast Asia

Item	PC1	PC2	PC3
Contribution ratio	0.59	0.16	0.07
Cumulative contribution rate	0.59	0.75	0.82

PC = principal component.
Source: Calculated based on ADB Asia SME Monitor 2023 database.

5.3. South Asia

In South Asia, three factors were obtained from the P-PCA, but PC1 and PC2 trends were swapped (Figure 3, Table 7). PC1 makes the largest contribution to the variation in country data (59%), followed by PC2 (22%) and PC3 (8%), explaining 89% in total (Table 8). Factor loadings are sorted in descending order (Table A3.3).

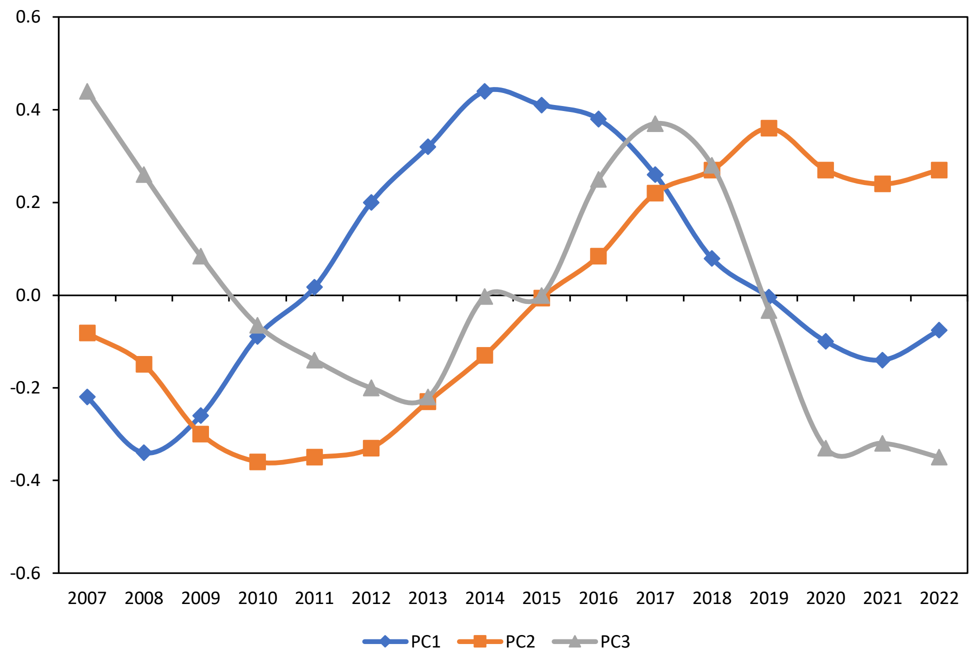
PC1-PC3 in South Asia generated different trend curves from “all countries” estimates (Figure 3). PC1 remains low until 2010, moves up until its 2014 peak, and then declines afterward. It indicates a fast recovery from the GFC (a year earlier than the trend in “all countries”) but was sensitive to the global economic slowdown and the pandemic. PC2 remains low until 2015, then rises through 2019, before declining until 2021. It indicates a slow recovery from the GFC and South Asia’s economic slowdown. The PC3 trend was somewhat reversed from the trend in “all countries.” It bottoms out in 2013, moves up to a peak in 2017, then drops afterward bottoming out in 2022. It indicates MSME development was sensitive to the region’s economic and political instability (e.g., India before the current administration started in 2014, and economic and political crises in Pakistan and Sri Lanka from around 2019) as well as the shock from the pandemic.

In PC1 (**fast recovery**), the negative curve until 2010 and after 2019 is explained by negative factor loadings denoted by nonperforming loans by banks, NBFIs, and for MSMEs (e.g., India, Bangladesh, Pakistan, and Sri Lanka). MSME loans in India and Pakistan were also identified as negative factors, suggesting that the delivery of low quality MSME loans with high levels of nonperforming loans in these countries likely impeded MSME development in the region. The positive curve peaking in 2014 is explained by positive factor loadings denoted by (i) NBFI loans and market capitalization (e.g., Sri Lanka, Pakistan, and Bangladesh) and (ii) number of MSME employees (e.g., India and Nepal). A recovery in nonbank and market-based finance, along with increased MSME jobs, likely supported a smooth shift back to growth. But the MSME funding environment was likely sensitive to economic and political environment changes, especially after 2019.

In PC2 (**slow recovery**), the negative curve until 2015 is explained by negative factor loadings denoted by MSME loans and nonperforming loans by banks, NBFIs, and for MSMEs in Pakistan and India, suggesting that increased MSME loans accompanying rising nonperforming loans in these countries likely made MSMEs recover slowly from the GFC and the region’s stagnant economic growth. The positive curve after 2016 is explained by positive factor loadings denoted by (i) loans by banks and for MSMEs (e.g., Bangladesh, Pakistan, Sri Lanka, and India), (ii) number of MSMEs (e.g., Nepal), and (iii) MSME output (e.g., Bangladesh and Pakistan). After 2016, improved bank lending and MSME loans along with an environment conducive to new small businesses and better productivity likely boosted MSME development.

In PC3 (**sensitive recovery**), the negative, downward trend during 2010–2013 and after 2019 (economic crises in Pakistan and Sri Lanka) is explained by negative factor loadings denoted by nonperforming loans by banks and NBFIs (e.g., Sri Lanka, Pakistan, and Bangladesh). The positive curve before the GFC and during 2016–2018 (linked to the new administration in India) is explained by positive factor loadings denoted by MSME loans and NBFI loans (e.g., Pakistan, Bangladesh, and India). PC3 was more affected by access to finance for MSME development. It suggests that high levels of nonperforming loans by banks and NBFIs likely led to a slowdown in MSME development. But once the MSME credit market and the nonbank finance industry expanded, MSME development quickly turned positive, although its growth pattern was likely highly sensitive to shocks, such as regional economic crises, political conditions, and the pandemic.

Figure 3: Time Series Plots of Estimated Principal Components—South Asia



PC = principal component.
Source: Calculated based on ADB Asia SME Monitor 2023 database.

Table 7: Time Series Plots of Estimated Principal Components—South Asia

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
PC1	-0.22	-0.34	-0.26	-0.09	0.02	0.20	0.32	0.44	0.41	0.38	0.26	0.08	-0.004	-0.10	-0.14	-0.08
PC2	-0.08	-0.15	-0.30	-0.36	-0.35	-0.33	-0.23	-0.13	-0.006	0.08	0.22	0.27	0.36	0.27	0.24	0.27
PC3	0.44	0.26	0.08	-0.07	-0.14	-0.20	-0.22	-0.003	-0.001	0.25	0.37	0.28	-0.03	-0.33	-0.32	-0.35

PC = principal component.
Source: Calculated based on ADB Asia SME Monitor 2023 database.

Table 8: Contribution of Each Estimated Principal Component—South Asia

Item	PC1	PC2	PC3
Contribution ratio	0.59	0.22	0.08
Cumulative contribution rate	0.59	0.80	0.89

PC = principal component.
Source: Calculated based on ADB Asia SME Monitor 2023 database.

5.4. Central and West Asia

As in other regions, three factors were obtained in Central and West Asia from the P-PCA (Figure 4, Table 9). PC1 makes the largest contribution to the variation in country data (63%), followed by PC2 (15%) and PC3 (8%), explaining 85% in total (Table 10). Factor loadings are sorted in descending order (Table A3.4).

PC1-PC3 show similar trends on MSME development as those in "all countries" (Figure 4). PC1 remains low until 2014, rising in 2015–2018 before slowing afterwards, indicating a slow recovery from the GFC. PC2 remains low until 2010, rises to a 2015 peak, and then declines with a negative curve after 2019, indicating a relatively rapid recovery from the GFC but sensitive to shocks like the RFC and the pandemic. PC3 bottoms out in 2009 (GFC), peaks in 2011, then drops until 2016 (RFC). It rises afterward with erratic movement during the pandemic, suggesting a quick but very sensitive recovery from shocks.

In PC1 (**slow recovery**), the negative curve until 2014 is explained by negative factor loadings denoted by (i) nonperforming bank loans (e.g., Kazakhstan, Azerbaijan, and Uzbekistan) and (ii) MSME loans (e.g., Kazakhstan, Tajikistan, and Georgia). The positive curve after 2015 is explained by positive factor loadings denoted by (i) bank loans and those for MSMEs (e.g., the Kyrgyz Republic, Armenia, and Georgia), (ii) number of MSMEs (e.g., Georgia, the Kyrgyz Republic, and Kazakhstan), and (iii) number of MSME employees (e.g., Azerbaijan, Kazakhstan, and Georgia). PC1 suggests that increased MSME lending with high levels of nonperforming loans in countries such as Kazakhstan likely slowed the recovery from the GFC. But improved bank lending likely supported creating new MSMEs and jobs in countries such as Georgia, helping them shift to growth.

In PC2 (**fast recovery**), the negative curve until 2010 and after 2019 is explained by negative factor loadings denoted by nonperforming loans by banks, NBFIs, and for MSMEs (e.g., Uzbekistan, Georgia, and the Kyrgyz Republic). The negative curve largely reflected the trends in Uzbekistan. The positive curve during 2011–2018—peaking in 2015—is explained by positive factor loadings denoted by (i) number of MSME employees (e.g., Uzbekistan, Georgia, and Tajikistan), (ii) MSME output (e.g., Tajikistan, Uzbekistan, and Georgia), and (iii) MSME exports and/or imports (e.g., Uzbekistan and the Kyrgyz Republic). PC2 suggests that a high level of nonperforming loans likely impeded MSME development. Increased job creation, higher output, and internationalization of MSMEs likely helped drive MSME development.

In PC3 (**sensitive recovery**), the downward curve around the GFC and RFC is also explained by negative factor loadings denoted by nonperforming loans by banks, NBFIs, and for MSMEs (e.g., Tajikistan, Uzbekistan, and Armenia). The negative curve in the PC3 largely reflected the trends in Tajikistan. The positive curve during 2010–2014 and after 2017 is explained by positive factor loadings denoted by (i) MSME exports and/or imports (e.g., the Kyrgyz Republic and Georgia), (ii) MSME output (e.g., Kazakhstan, Azerbaijan, Georgia, and the Kyrgyz Republic), and (iii) NBFI loans (the Kyrgyz Republic). PC3 suggests that MSMEs felt the hard impacts from financial crises (GFC and RFC) with poor access to quality bank credit and NBFI loans, more pronounced in Tajikistan. However, higher MSME foreign trade, output, and improved access to NBFI loans likely encouraged MSME development, yet it remained very sensitive to shocks like the GFC, RFC, and the pandemic.