3

APPROACH AND METHODOLOGICAL NOTES

3.1 Exploration: Desk research and discussions with industry experts[7]

The first step in conducting the study was an observation of the field and issues departing from the terms of reference. The idea was to first obtain (preliminary) insight into research of skill issues, labour market development in Malaysia at large and Penang specifically, and existing studies carried out earlier in Malaysia and Penang. Next, we gain understanding from stakeholders on issues related to labour market situation, which helps to fine-tune the focus of this study. An extensive desk research was carried out, followed by meetings and discussions with stakeholders, resource persons, and industry experts. These revolved around Penang skill situation, trends, and labour market processes, as well as issues concerning the target group (higher-qualified or professional workers). Also, expectations with regard to future industry trends and their implications were gauged. Table 3.1 details tasks carried out in the preliminary phase.

Table 3.1: Detailed description of tasks in preliminary phase

Step	Task	Description
1	Perform desk research trends	Provided useful insights into current issues and likely major trends in industries relevant to Penang as well as into ensuing current (and future) skilled labour needs in Penang.
2	Perform desk research	Existing studies of skill gaps and shortages, international; existing reports of skill gaps and shortages in Malaysia.
3	Perform desk research: Study demand and supply framework	The framework of labour demand and supply was studied to understand the labour market theory, and to segment Penang labour force into different compartments based on firm size and origin of a firm.
4	Define target agencies & individuals for initial interviews with resources persons	Experts include industry leaders, industry consultants, policy makers, and academics who are continuously involved in the planning and analysis of human capital development in Penang in relation to desired/targeted evolution of sectors/ industries/branches. Targeted persons were well-informed representatives of agencies and associations including the Department of Skills Development (Northern region), Penang Career Assistance and Talent Centre (Penang CAT Centre), TalentCorp Malaysia, Labour Recruitment and Employment Agencies, Federation of Malaysian Manufacturers (FMM Northern Branch), Northern Corridor Implementation Authority (NCIA) and others. In part, experts and stakeholders are identified on the basis of researchers' past experiences and consultations with resource persons. A number of criteria are used to shortlist experts. These include working experiences; knowledge; his or her contributions towards a specific industry group; seniority of position or rank in industry associations.
5	Conduct interviews with resource persons	Information gathering was conducted through personal interviews held with a select number of local experts and agencies. Information gathered was primarily qualitative. Insights from individual experts were anonymously shared in the focus group discussions to collect responses from industry players.

[7] This chapter builds on the Interim Report of this study, July 2016; and on Terhorst, J. & T. Verbraeken (2016). *Making the Transition into a High-Income Economy: the Penang Case*. Master Thesis Economic Geography, Utrecht University, The Netherlands.

Based on (the results of) the observation, the research design, topics, and knowledge aspects were defined. These are outlined in the remainder of this chapter. We end the discussion by outlining some major obstacles that surface in the course of implementation of the study and the limitations following from these.

3.2 Selection of priority industries

The priority industry is defined based on its economic importance to the growth of Penang, and the industry focus is presented in Figure 3.1. The guiding principles further impinge on industry upgrading and diversification shaping Penang's Next Economy (that is, transition from a near exclusive manufacturing to a manufacturing and services economy).

3.3 Approach to the study and information gathering methods

The identification of indicators at macro and micro (meaning firms and workers) levels for defining and measuring skill shortages and gaps, has led to the study being divided into three components. Figure 3.2 depicts these three components that are distinct, and therefore have been carried out independently. The first component deals with the demand and supply side characteristics of the relevant part of the labour market, as indicated by relevant vacancies and output from educational institutions. At meso- and micro-levels, the second component examines the skill needs and skill gaps of firms and their industries, while the third component scrutinises the characteristics and drivers of high-skilled labour mobility. The first component also entails the identification of future trends in the Penang economy, and ensuing skill requirements through information gathering from experts and relevant individuals/agencies.

The approach is novel in three ways: first, research in the Penang context has focused on a single component rather than all three simultaneously in an integrated manner. Second, this is the first study to explicitly consider worker mobility in the context of skill shortages.

Figure 3.1: Industry focus of this study

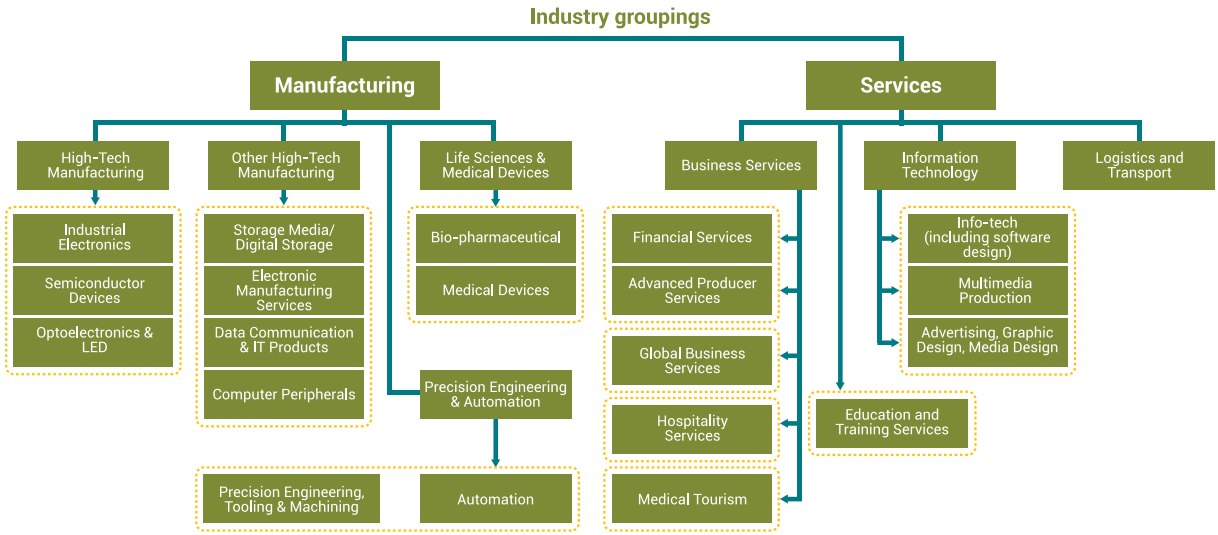
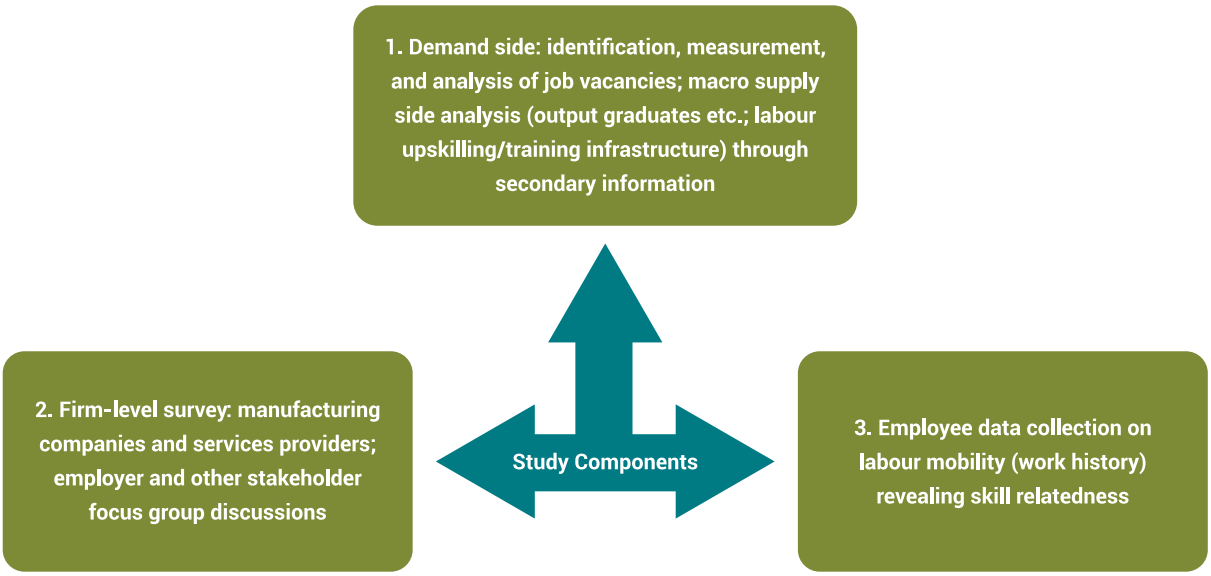


Figure 3.2: Set-up of the study



Third, besides the emulation of common information gathering methods employed in studies in this field, digital information sources that are still in the category of unconventional or experimental have been tapped. These will be explained as we discuss the work flow. The collection and analysis of data have proven to be particularly challenging as far as data processing and information analysis are concerned.

3.4 Methodological steps

The three components use mixed – both quantitative- and qualitative-oriented methods. Multiple research design is a means for understanding skill shortages from diverse viewpoints. The mixed method research design with respect to each component is explained in Figure 3.3, which depicts the elaboration of each component in methodological steps.

3.4.1 Component 1: Job vacancies, demand- and supply-side analysis

Component 1 identifies hard-to-fill vacancies and examines graduate output to meet the skill needs. Figure 3.3 illustrates the steps to administering this component, which includes defining indicators and information sources, mining job vacancy data and validating data through focus group discussions and conducting survey among companies with multiple vacancies. The specific methodological tasks of this component are detailed in Table 3.2.

Figure 3.3: Components of the study and elaboration in steps

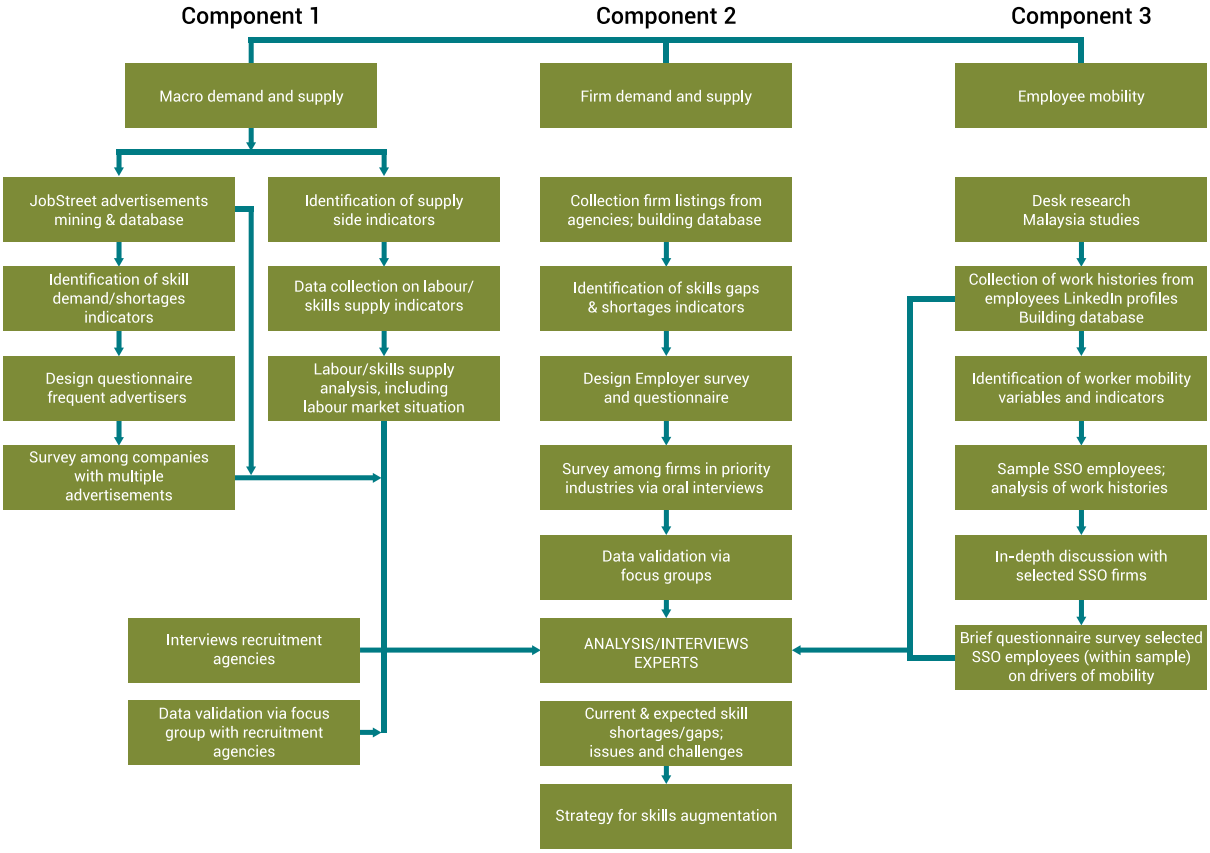


Table 3.2: Detailed description of tasks in Component 1

Step	Task	Description
1	Identification of demand and supply side indicators	See Section (a).
2	Gather secondary data on supply side	As part of this component of demand and supply side, information has been assembled from a range of sources that include the Department of Statistics (DOSM), Ministry of Higher Education (MOHE), training centres, and others. Data were also gathered on the existing upskilling infrastructure (especially private training providers) in Penang. See Section (a) for further details.
3	Define information sources on vacancies	Making an inventory of job vacancies and tracking them over a span of six months enable analysis of vacancy characteristics, including when and how they are filled. A number of criteria have been applied to assemble job vacancy listings. While the primary criterion is target-industry, vacancies in the high-skilled category weigh strong in the analysis. See Section (b).

Table 3.2: Detailed description of tasks in Component 1

4	Mine job vacancies	The main channels used to mine job vacancies are job portals and employment/ recruitment agencies such as Jobstreet.com, Kelly Services, and Penang CAT Centre. However, Jobstreet.com is the largest online job portal in the region and in Malaysia. Therefore, these vacancies have been mined over a span of about six months from the last week of December 2015 to June 2016.
5	Build a database of vacancies and skill requirements	A substantial number of (new) job advertisements appear online every month in our target industries. Within Penang region, approximately 2,000 advertisements are placed on a fortnightly basis (including new and re-postings vacancies). The nature of job postings has been observed, which consist of vacancies outlining job descriptions and skill requirements. These are posted online for a month, after which will be renewed by the advertising company (if the vacancies have not been filled yet); while some companies continuously advertise for the same position(s).
6	Conduct survey among companies with multiple vacancies	To gauge responses on advertised vacancies, selected employers have been approached to provide information on the success rate of filling the vacancies. Of particular interest were companies that frequently advertised through job portal. See Section (c).
7	Interview recruitment agencies	Insights from professional recruitment agencies have been obtained from face-to-face interviews. The recruitment agencies serve several roles in the labour market, and they have the best knowledge about the demand and supply of high-qualified skilled labour and the range of skills. Interviews have been conducted with selected recruitment agencies that provide job placement services for executive positions and higher, including expatriate staffing.
8	Validate data via focus group discussion with recruitment agencies	To validate analysis of demand and supply information, a focus group discussion with professional recruitment agencies has been organised and conducted. For the purpose of this discussion, as well as focus group meetings with other groups in other study components, a protocol has been developed. Such focus group discussion has given the opportunity to discuss: a) the relevance and comprehensiveness of the Critical Occupation List (COL) approach employed by TalentCorp[8], and the relevance of COL or Critical Skills List (CSL) pertaining to labour and skill needs of the industry; b) the contents of a Penang-specific COL/CSL; and c) the correspondence of critical occupations and hard-to-fill positions.
9	Cleaning and analysis of job vacancy database	See Section (d) below and Technical Report: Annex 7
10	Labour/skills supply analysis, including labour market situation	A major focus here is – the patterns of – recruitment difficulties, through the identification and skill characteristics of 'persistent' vacancies, and the occurrence of such vacancies in the industries and firms.

a) Macro demand and supply side indicators

In general, macroeconomic indicators for labour and skill shortages mainly consist of quantitative data provided by national agencies. These are to identify employment situation and assess labour market conditions from demand and supply perspectives. Table 3.3 provides an overview of all indicators used to measure the variables of labour and skill shortages at macro level. (See also Chapter 2, Tables 2.1; Technical Report: Annex 4)

The following section will briefly outline how the different indicators have been used to measure the variables. The first two indicators are combined under

[8] See TalentCorp (2015); Ilmia and TalentCorp (2016)

the heading *'(un)Employment'*, indicators three and four are captured by the heading *'Qualifications of output (supply) of educational institutions vs. demand, upskilling infrastructure and retention'*. Finally, the demand side indicators five and six have been amalgamated under the heading of 'vacancies'.

(un)Employment

The employment and unemployment rates are often used as indicators for labour and skill shortages. Unemployment can be an indication of qualitative imbalances in supply and demand, which could either be oversupply, undersupply, or mismatch[9]. A high unemployment rate could indicate excess supply of high-, intermediate or low-qualified labour. It can merely signify the presence of additional stock of labour. However, as noted in Chapter 2, an unusually high unemployment rate can result from employability issues associated with skill deficiencies, rather than indicating labour surplus[10]. Thus, Migration Advisory Committee (MAC, 2008b) considers unemployment rate also as an indirect indicator for skill shortages[11]. One has to note, however, that this appears contingent on overall labour market conditions. It may be less suitable in case of a tight labour market.

Unemployment could occur in certain occupations where specific educational qualifications are required. When actual skill sets are not in line with these required qualifications, employability issues occur. As a result of changes in firms' production technology, demand for workers, particularly occupations associated with high levels of education (e.g. professionals and managers) increases. When supply primarily consists of people with low levels of education, a demand-supply mismatch of skills at different educational levels will occur. Eventually, this leads to redundancy and high unemployment among low- and intermediate-educated workers. At the same time, the overall economy experiences shortages of highly educated/ skilled workers[12].

Table 3.3: Macro level variables, indicators and sources

Scale	Variable(s)	Indicators	Sources
Macro	Labour and skill shortages	**Supply side indicators**	
		1. Employed persons	• Department of Statistics Malaysia
		2. Unemployment rate	• Department of Statistics Malaysia
		3. Entrant into labour market (graduates): by field of study	• Ministry of Higher Education • Interviews/focus groups
		4. Extensiveness of labour Training/skilling infrastructure	• Interviews/focus groups • Inventory institutions/agencies • Employer survey
		Demand side indicators	
		5. Vacancies	• Department of Statistics Malaysia
		6. Vacancy fill rates and hard-to-fill vacancies	• Vacancy data and analysis, obtained from agencies and employer survey

9 See European Centre for the Development of Vocational Training (Cedefop, 2012a and 2012b)
10 See Shah and Burke (2003)
11 See Migration Advisory Committee (MAC, 2008a and 2008b)
12 See Kahn (2015)

The downside of using unemployment rate as an indicator for skill shortages is that unemployed persons could be voluntarily out of work (which is not due to unavailability of jobs in the economy). This leads to an overestimation of supply[13]. Also, while looking for a (new) job, a qualified person may temporarily be employed in lower-qualified jobs. This in contrast leads to underestimation of supply of labour that possesses a specific skill-set. However, this indicator is relevant to assess potentially existing skill shortages[14].

Qualifications of labour supply vs. demand, upskilling infrastructure and retraining

Supply of labour with specific educational attainment is often related to skill shortages. Educational attainment indicates employability of graduates in the labour market. If graduates meet the required skills and qualifications from employers, firms are more likely to fill the vacancies. Furthermore, many studies have highlighted discrepancies between the curriculum in tertiary education institutions and the competences required by employers. Also, competences expected on the basis of educational qualification and the skills job seekers can offer are largely lacking[15].

Although the balance between working skills and formal education must be timely mitigated, the education industry is often slow in responding to observed skill shortages. This is exacerbated by the different levels of specialisation and the rapidly changing skill-sets required by employers (MAC, 2008b). Even if educational institutions respond quickly, there always remains a time lag before the changes have impact.

An efficient approach is to detect demand for new qualifications early. The slow response from actors and educational institutions may amplify skill shortages and increase graduate unemployment. A better co-operation and interaction between corporate and educational/training institutions is crucial to alleviate these concerns.[16]

Vacancies

Another indicator that detects imbalances in the labour market is the number of vacancies[17]. When the number of vacancies increases, demand outstrips supply indicating an increase in the number of unfilled vacancies. This is more likely to happen to job positions and/or occupations in a particular industry or range of industries than within the same or other industries. Therefore, the breakdown of job vacancies into functional and/or occupational and/or different industries is necessary to estimate skill shortages.

The number of job vacancies is equivalent to the number of employees that make jobs vacant and the number of jobs created. In general, expansion of firms leads to new jobs. Also, when a new company decides to establish in a region, new jobs will be created. The expansion and set-up of new firms may cause skill shortages when the supply of skills in that particular region does not match the requirements of new and existing firms.

When both unemployment and vacancy numbers are available the ratio (V/U) can be used to study the relationship between employment and vacancies at the aggregate level.

Aside from the growing number of vacancies and vacancy ratios, qualitative aspects of vacancies are more of interest and relevance to address skill shortages. The average duration or search time of vacancies relate to the length of period taken for a job vacancy to be filled. Many studies make use of hard-to-fill vacancies as an indicator of skill shortages. When the duration of vacancies increases, they are referred as hard-to-fill vacancies (HTFVs), which implies that it takes a longer duration (than usual) to hire a suitable worker[18]. HTFVs or recruitment difficulties can, however, occur for several other reasons than skill shortages, such as the conditions of work offered (e.g. wages or work hours) and employers' reputation[19]. Furthermore, instead of indicating a skill shortage that needs to be addressed through market intervention, hard-to-fill vacancies may simply indicate a high turnover within the particular occupation[20]. In addition, in the case of vacancies for non-manual labour such as healthcare workers, sales workers and others, search time is longer as suitable candidates are more frequently drawn from existing market pool[21].

13 See Shah and Burke (2003)
14 See Veneri (1999)
15 See Teijeiro et al. (2013); Froy (2013)
16 See Rahman et al. (2010); JPMorgan Chase & Co. (2014); Pauw et al. (2008)
17 See MAC (2008b); Shah & Burke (2003)
18 See Haskel & Martin (1993, 2001)
19 See MAC (2008b)
20 See MAC (2008b)
21 See Andrews et al. (2008)

b) Information and data sources

General labour market information

The secondary data on macro level was obtained from the following sources: Department of Statistics Malaysia (DOSM), Ministry of Human Resources Malaysia (MOHR) and Ministry of Higher Education Malaysia (MOHE). These sources were used to obtain demand and supply side statistics of the labour market. Statistics on demand in the labour market include vacancies to assess labour and skill needs. On the other side, supply statistics consistof (un)employment numbers, share of graduates by field of study, and the extensiveness of labour training and skilling infrastructure. The Social Statistical Bulletin published by DOSM records the number of new job seekers for males and females.

We collected the number of new job seekers over the past ten years (2005–2014) so that the trend over time is studied. This was then compared with the number of entrants into the labour market. The best proxy for entrants is the number of fresh graduates in a range of study fields. These data have been obtained from various publications by MOHE, including the Graduate Tracer Study Report. Second, ILMIA (Institute for Labour Market Information and Analysis; part of the MOHR) provides information on the number of employed persons by industry and occupation, and labour force statistics (including labour force participation rate and unemployment rate). Such data are also made available by DOSM on special request. The data are, whenever possible, gathered on state-level, otherwise national data are used as a proxy. These statistics offer a picture of the current trends and conditions in the Penang's labour market. It involves statistics on labour force characteristics (such as participation rate), (un) employment numbers and characteristics), vacancies and graduates by field of study.

Digital Portal Vacancies

Besides statistical sources, data on vacancies were obtained from JobStreet. This portal is commonly used by companies to advertise vacancies. For the Penang region, approximately 2,000 positions are advertised every two weeks (both new and repostings). In Annexes 5–7 of Technical Report, collection of advertisements, relevant criteria and the building of a database are explained.

It should be noted that vacancies for higher-qualified positions are not fully captured by the vacancy advertisements. For strategic or key positions, employers prefer to engage employment/recruitment agencies rather than placing an advertisement.

All advertisements have been mined bi-weekly for a period of six months, from 30 December 2015 to 30 June 2016. Since complete datasets could not be made available by recruitment firms, job advertisements have been manually mined from JobStreet.com using sectors and job specialisations corresponding to target-industry branches. The job portal covers most industries and occupational categories in the labour market. Available search engine and classification by online job portal have been used for initial filtration. However, we have devised our own scheme to reclassify. Following initial mining, the 'behaviour' of postings has been monitored. Box 3.1 explains some observations on vacancies mining.

Box 3.1: General observations on vacancies mining

A number of scenarios are possible with respect to online advertising for vacancies. First, the vacancy is filled within the month, thus it will not be re-advertised. Second, the vacancy is not filled within the advertising period, and the company renews the posting, hence the job is advertised again in the following month.

In principle, mining every two weeks allows – through comparative analysis – identification of not only vacancies that remain open for a substantial period, but also those that are filled rather fast. However, through special accounts companies can continue advertising positions for an extended (even indefinite) period, as demand and supply dictate (e.g. advertisement results in successful recruitment but the advertisement stays open as a number of identical positions are available). It could also be the case that the vacancy is no longer advertised at the portal because it is withdrawn or has been filled internally. Such and other company behaviours limit the usefulness of length of advertising period as an indicator of hard-to-fill vacancies. Some additional caveats should be noted. Some companies engage in the practice of posting 'false' job vacancies to test/assess the current market condition. Also, employers may advertise multiple vacancies without having a certain number of positions in mind. The job portal does not trace whether and when a vacancy is being filled, or practices such as those described above.

Special software has been used to directly create Excel

files from vacancy information. This substantially helped in building a database of advertisements/ vacancies. The database comprised over 20,000 entries. Skill requirements are among the variables recorded for each advertised position (see Technical Report: Annex 7).

Interviews/focus groups recruitment agencies and educational institutions

Various primary and secondary sources have been used to obtain macro level information and data. Primary sources included semi-structured interviews and focus group discussions with recruitment agencies and educational institutions (see Technical Report: Annex 14 for topic lists for recruitment firms, life sciences and educational institutions). Additionally, semi-structured interviews with recruitment firms, training institutions and industry stakeholders were used as primary source of data (see Appendix L).

To obtain more in-depth qualitative information and insights into issues at hand, focus group discussions have been held with selected recruitment agencies and educational institutions. Focus group meetings also presented an opportunity to compile a Penang-specific Critical Occupation List (COL)/Critical Skills List (CSL), covering the demand side. This COL/CSL can be validated with the vacancies database compiled. A recruitment survey was also carried out to gather participants' opinions on skill demand, skill shortage, and challenges in recruitment of high-qualified employees. It is worth noting that most of the small recruitment firms did not attend the meeting. Attempts have been made to approach all firms who were absent from the focus group meeting, with a request to fill out the Penang-specific COL/CSL. This was done via email and by visiting the office. Only two firms responded positively.

c) Survey companies with multiple advertisements

To find out what happened to the different vacancies that were advertised, a survey was conducted by sending out a brief online questionnaire to selected advertising firms. In view of the large number of vacancies posted, it was decided against following-up a broad selection of postings every 14 days. Instead, to analyse the skills required by companies, responses, and filling of vacancies, as well as to trail repeat job postings, it was decided that focus would be made on companies that regularly post a substantial number of vacancies. An online survey method was opted because it was less time consuming than sending out letters and mailing questionnaires. Furthermore, it is more convenient for firms, likely producing a higher response rate. The questionnaire contained questions about the causes of vacant positions, how they were filled (by (un) qualified workers), and whether companies experience recruitment difficulties (see Technical Report: Annex 12, Section C). The survey schedule was incorporated into the employer survey as many firms in the database concern companies that advertise for vacancies.

Firms that were approached for this survey were selected on the basis of the number of vacancy advertisements in the given time period. A high threshold was used since a majority (about 70%) of companies had advertised between one to five vacancies. For these companies, hardly any useful results were obtained. On the other hand, the companies posting a large number of job advertisements at the same time 'took' the large majority of advertisements/vacancies. All firms selected had been posting above the average number of 11 job advertisements. This resulted in the inclusion of 100 firms that were not listed in the company database.

A major obstacle to administering the vacancy survey schedule was the lack of company (specifically HR manager) contact information as this is not included in online job advertisements. Contact information had to be obtained from various sources and/or by making phone calls. This proved to be a tedious and slow process. For this particular survey, contacts of 130 firms were collected out of the 195 firms selected. The online survey was sent to these 130 firms, and 33 firms returned the survey schedule, producing a response rate of about 25%.

d) Cleaning, organising and analysing data

Primary data

Interviews with resource persons and focus group

discussions have been recorded (subject to approval) and notes taken. The needed information has been extracted from these.

Secondary statistical data

All statistics obtained from institutions have been transferred to Microsoft Excel. This software allows conversion of data into charts and tables.

Vacancy advertisements database

A range of procedures have been employed in regard to processing and analysis of the vacancy database. An elaborate outline of these procedures is given in the Technical Report: Annex 7, and they are summarised as follows.

1. A filter was applied to extract advertisements/ vacancies by/in companies/industries within the scope of this study (see selected key industries).
2. Next, job titles that recurred across multiple subsequent mining were deleted, as these are assumed to be equivalent to unique job openings.
3. The number of unique vacancies was estimated by filtering all identical combinations of company names, job positions, and job titles. Using these criteria, there are 4,455 unique vacancies, with an average of 2.58 advertisements per vacancy and most unique vacancies are being advertised thrice.
4. By applying mining thresholds and order, high-demand positions and 'persistent' vacancies have been determined among all unique vacancies. Persistent in many cases is considered to be hard-to-fill.
5. To reduce job titles and functions of high-demand positions and persistent vacancies to a manageable number (range), classification was applied.
6. Assigning key words was also applied to arrive at a manageable listing of most frequently asked job titles and functions. In addition, to keep the number of vacancies manageable, a selection procedure was applied to companies advertising multiple vacancies.

3.4.2 Component 2: Industry and firm-level analysis

Component 2 consists of an industry- and firm-level analysis of the skill situation in Penang. For each priority industry branch, information on skill shortages and skill gaps is obtained from local and foreign firms; large, medium and small firms respectively. Figure 3.3 shows a number of steps to accomplish the objective of this component. These include building a firm database, identifying skill gaps and shortages indicators, survey and questionnaire design, up to validating data through focus groups discussion and data analysis. Methodologically, each step is further described in Table 3.4.

Table 3.4: Detailed description of tasks in Component 2

Step	Task	Description
1	Build firm database	With the use of sources from state and federal government agencies, industry associations and resource persons, a database of companies in our priority industries has been compiled. This database has a total of about 1,200 firms. See Section (a) below.
2	Identify skill gap/shortage indicators	Skill gap/shortage indicators have been obtained from literature and studies carried out elsewhere. Indicators have been adapted to the local context. Skill issues may vary among firms according to origin (local or foreign), size, and industry branch. See Section (b) below.
3	Develop survey design	As part of this study, an employer survey has been conducted to collect firm-level data pertaining to demand and supply of skills, and skill utilisation. Decisions in regard to actually conducting the survey were taken and revised as this component evolved. See Section (c) below.

Table 3.4: Detailed description of tasks in Component 2

4	Develop questionnaire	A questionnaire has been developed for the employer survey. It follows questionnaires designed for similar surveys conducted in UK by the United Kingdom Commission for Employment and Skills (UKCES, 2013), in European countries by the European Centre for the Development of Vocational Training (Cedefop, 2013), and by World Bank[22]. Also, questionnaires employed in consultant studies carried out earlier in Malaysia have been considered[23]. For questionnaire: Refer to Technical Report: Annex12, and Section (c) below.
5	Select firms	Firms whose businesses belong to the priority industry branches constitute the population of the company survey. Criteria for firm selection for participation in the survey were devised as this component evolved. See Section (c) below.
6	Administer questionnaire	The questionnaire has been administered among selected firms from the company database. Due to poor response to the online survey, face-to-face interviews have been conducted with selected firms. In the end, interviewed firms were representative of the population in terms of numbers, industry, and firm distributions. See Section (c) below.
7	Validate data via focus group discussion for respective industries	Focus group discussions were envisaged with groups of employers representing the key industries. Focus group discussions with specific industries were meant to validate the survey information gathered. For the purpose of this discussion, a protocol similar to the one used for the employment/recruitment agencies has been employed. These focus group discussions also gave the opportunity to discuss: a) The relevance and comprehensiveness of the Critical Occupation List (COL) approach employed by TalentCorp, and the relevance of COL or CSL to express skilled labour requirements of and availability to industry; b) The contents of a Penang-specific COL/CSL; and c) The correspondence of critical occupations and hard-to-fill positions.
8	Collect upskilling/ training infrastructure	As the company survey also covers human capital training, an inventory of the current training infrastructure in Penang has been made. A range of sources have been used for this purpose.
9	Data analysis	See below

[22] United Kingdom Commission for Employment and Skills (UKCES, 2013); Employer Skills Survey 2013. Technical Report. London: UKCES; European Centre for the Development of Vocational Training (Cedefop) (2013). User guide to developing an employer survey on skill needs. Publication Office of the European Union, Luxembourg; Gaëlle et al. (2014)

[23] ILMIA/UPM (2016); IPSOS (2012, 2014a, 2014b); PwC (2013a, 2013b); World Bank and ILMIA (2014)

a) Building a company database

An attempt is made to compile a relevant/valid company database through company listings acquired from local authorities and federal government agencies. These include MBPP, MPSP, SSM and MIDA. MPSP and MBPP provide lists of firms, which have applied for licenses to operate businesses in Seberang Perai and Penang Island. There prove to be significant drawbacks and deficiencies in the listings provided by individual agencies.

Despite the fact that the lists provided by both local councils could be complete, MBPP's list was of little use due to lack of convertibility in a format necessary for compiling our targeted database. The main drawback of the firm listing furnished by MPSP pertained to the business activity not being classified according to the standard schema. However, it was still used as a reference in developing our company database, and some reclassification was done using the Malaysian Standard Industrial Classification (MSIC) 2008. General issues are the incongruities of industry coding between the listings obtained from the various agencies, as well as apparent lack of comprehensiveness.

SSM's firm listing appears to have a suitable coverage: it captures existing companies – incorporated in Penang – registered with the Registrar of Companies (ROC). SSM uses MSIC (2008) to classify company's business activity. A major issue that surfaces when we go through this firm listing is the manifold mismatches between industry coding and actual nature of business. Due to errors and inconsistencies in allocation of industry code, individual codes produce a heterogeneous set of firms, many of which are actually not part of the target industry. Also, firms in a target industry are found under non-expected and multiple industry codes but it is not possible to immediately detect which ones. A fuzzy coding implies the risk of missing a substantial number of firms in a target industry.

We manage to develop a procedure to resolve this issue, and establish the codes set for our target industries[24]. Although this procedure is feasible, we decide to employ a more direct approach by approaching specific industry associations and consulting websites such as booking.com, agoda.com and so on. Table 3.5 lists agencies that provide company lists.

b) Meso- and micro-level skill shortage and gap indicators

At the meso-level, the indicators shown in Table 3.6 have been used.

At the micro-level, the indicators shown in Table 3.7 have been used.

Table 3.5: Institutions that provide company lists

Industry organisations	Local authorities	Recruitment agencies
• Federation of Malaysian Manufactures (FMM) • Malaysia Biotech Corp • Penang Foundry & Engineering Industries Association (PENFEIA) • Association of private hospitals of Malaysia (APHA) • Malaysia International Chamber of Commerce and Industry (MICCI)	• InvestPenang • TalentCorp • Malaysia Investment Development Authority (MIDA)	• Kelly Services • EPS Consultancy

[24] This procedure involved using the MIDA database as entry, additional collection of firm listings from associations, clubs, consortia, and so on (FMM, Chambers of Commerce, Penang Industrial Clusters and resource persons), for codes identification, cross-checking, and validation purposes.

Table 3.6: Meso-level variables, indicators and sources

Scale	Variable(s)	Indicators	Sources
Sectors/industries	Skill shortagess	1. Hard-to-fill-vacancies 2. Most frequently asked job titles	• Vacancy data analysis • Focus groups • Employer survey
	Skill deficiencies and gaps	3. Training 4. Skills that need the most improvement	• Focus groups • Employer survey

Table 3.7: Micro-level variables, indicators and sources[25]

Micro	Variables	Indicators	Sources
Firms	Skill shortages	1. Hard-to-fill vacancies 2. Positions not filled 3. Length of time to fill vacancies 4. % of applicants fully meet requirements 5. Labour poaching 6. Firms' recruitment standards 7. Labour turnover 8. Under hiring	• Employer survey • Focus groups
	Skill gaps	1. Skill proficiency level 2. % of fully skilled employees 3. Skills that need the most improvement 4. Preparedness of employees' new tasks 5. Labour turnover 6. Training	• Employer survey • Focus groups

Recruitment standards include educational qualifications, wage limits, fringe benefits and language level, to name a few. In the search for skills, employers will mostly recruit internally before they outsource in the external labour market. In a tight labour market characterised by high levels of demand, firms' recruitment standards might decrease compared with markets with high levels of supply. In order to find a suitable candidate who possesses the right skills (language proficiency and educational qualifications), employers might oblige by offering higher wages (premium), interesting fringe benefits, or lower qualification requirements. Firms that are willing to lower standards are expected to face skill gaps rather than shortages. Employers offering high(er) wages within firms can be an indication of skill gaps[26]. To gauge firm recruitment standards, employers have been asked if they provide special incentives to hire a candidate who possesses all the required skills and is able to choose among offers from different employers.

[25] See Andrews et al. (2008); Combes & Duranton (2003); Sheldon & Li (2013); Stevens (1996)
[26] See Schlicht (2002)

c) Survey design, questionnaire, company selection and administering questionnaires

An *online survey was* originally envisaged. However, given the limited information available on individual companies, additional information has to be collected for such a method to be employed. Companies were contacted by phone to verify their existence, their industry branch, persons-in-charge of personnel matters, and their email addresses and contact numbers.

An introduction letter explaining the motivation of this study was mailed to all firms in the database. This is to ensure that the target companies are aware of this study; the letter also emphasised the benefits of participating in the survey. Unfortunately, a number of letters were returned because of incorrect address, relocation and company closure.

Subsequently, after verifying the telephone numbers, companies were contacted to collect the required information. It is important to note that a substantial number of telephone numbers were either not in service or incorrect; also, when contacted, a number of respondents (16% of companies, including Financial Services companies where all HR operations are headquartered in Kuala Lumpur) declined to provide information as they were unwilling to participate in the survey.

The questionnaire consists of seven sections[27]. The first section has questions on company particulars such as name, business activity and origin. The second section has questions pertaining to current employment in the firm. The third section covers questions on skill requirements of current workers and skill gaps. This is followed by the fourth section with questions on the impact of and remedial measures to skill gaps. Questions on labour recruitment, skill needs and shortages constitute the fifth section. Questions on the consequences and responses to skill shortage and firm's views about the local labour market, skill situation, and the way skill issues should be addressed are covered in the last two sections of the questionnaire. Naturally, question-and-answer categories (closed questions) are adjusted to the Penang context.

The earlier interviews with recruitment companies and agencies also served the purpose of allowing validation of the questionnaire in regard to operationalisation and measurement of variables, indicators, and set-up. To test the schedule, a pilot survey among a small number of companies from different industry branches was carried out. Three companies from Medical Devices, Hospitality Services and Professional Services participated in the oral pilot interviews.

The questionnaire was then adapted based on the feedback and responses from the pilot survey. As the length of questionnaire can be a factor influencing survey participation, care was taken to limit the length of the survey to reduce non-responsive outcome due to the amount of information asked.

Out of the 1,200 companies in the database, less than half of the attempt to obtain contact information was successful. All companies with contact information were initially included in online surveying.

These companies were sent a second letter reiterating the nature and importance of the study, the importance of participation and the benefits for the company. This was followed up with information on the link to access the online survey.

After the initial limited response to the online survey, many calls were made to these companies, specifically the persons-in-charge of HR to explain the objectives of the online survey, the importance of participation, and the deadline for submission.

As response remained poor, several strategic changes have been undertaken to increase the responses. These include:-

1. The survey schedule was shortened, making it less time consuming.
2. The shorter survey schedule was emailed to all companies for which email address was available. In an accompanying letter, companies were again

[27] See Schwalje (2012)

invited to participate.

3. Selected companies were approached directly with a request for an oral interview with the top management through contacts shared by industry associations. Some recruitment firms have also been helpful in providing contact information of human resources managers in relevant companies.
4. Due to the low rate of responses, additional assistance from a number of industry associations was sought for distribution of the questionnaire among their member companies. Leaders of various industry associations have been contacted to urge members in their regular committee meetings to return the questionnaire. Associations that have assisted in this are FMM, PSDC and InvestPenang.
5. In addition, calls were constantly made to remind invited companies to return the questionnaire and to make appointment for oral interviews.

Finally, a satisfactory response rate was achieved with a total of 92 responses in about six months of data collection.

d) Focus group discussions

Unfortunately, a number of envisaged focus group discussions, it proved rather impossible to bring together sufficient participants, as a result of which a number actually could not proceed.

3.4.3 Component 3: Employee analysis on mobility behaviour

The third component examines labour mobility at employee level. From this, patterns of skill-relatedness can be deduced[28]. For a number of reasons to be clarified below, research was exploratory and experimental rather than encompassing comprehensive analysis. Figure 3.3 demonstrates six of the steps comprising this study component. The detailed description of each step is given in Table 3.8.

Table 3.8: Detailed description of tasks in Component 3

Step	Task	Description
1	Desk research	Through various channels, studies on mobility of higher-qualified labour were searched. The focus was on Malaysia. Beyond the aspect of brain drain, little studies can be found. While several were obtained, it is clear that labour mobility is still an under-researched area in Malaysia.
2	Collect employee work histories Define information source	To examine labour mobility across industry branches and occupations in Penang, as well as skill-relatedness, data have been collected on work/job history of individual employees. LinkedIn profiles constitute the main source of information. A database of work histories have been constructed by extraction and downloading LinkedIn profiles supplemented by information from other sources, and translating these into a MS Excel file. See Section (a) below.
3	Identify drivers, variables and indicators for mobility	Based on the drivers, relevant variables and operational indicators concerning labour mobility have been determined.
4.	Sample selected employees for analysis	For practical reasons, a subset of employee work histories was sampled from the LinkedIn profiles for one single destination industry.

[28] See Otto & Weyh (2014)

5	In-depth discussions firms	In-depth interviews were conducted with selected firms to uncover more about their labour recruitment strategies and sources (specifically the role of the secondary supply), preferences as to background of workers, labour turnover and HR management practices, and other mobility aspects.
6	Develop and administer a brief questionnaire among a selection of employees in the earlier sample	A brief online questionnaire was implemented to gather additional information on employees' work histories and their behaviours in job change. The questionnaire has been developed using Google Form. As to selection of employees, persons in the earlier LinkedIn profiles sample were contacted for whom a contact email address was available. Thus, the main method used to administer the questionnaire was dissemination by email.
7	Build a questionnaire data file	Completed questionnaires provide the input for a data file of individuals. See Section (b) below.
8	Data analysis and presentation	See Section (c) below.

a) Database of work histories

Jobs Malaysia was approached to obtain CVs of job seekers as it is a public domain kept by the Ministry of Human Resources. It holds open access CVs of job seekers. However, it was decided not to use this source. It proved to be hard to access; furthermore, all information is administered by the headquarters situated in Kuala Lumpur, rendering compilation of CVs difficult. While a search made at the portal of Jobs Malaysia revealed nearly 9,000 CVs stating Penang as the (preferred) place of work, compilation of CVs would be much simpler if Jobs Malaysia is able to provide a database in a ready-to-process format. Thus, LinkedIn was the most feasible source to collect CVs or work history records of employees in Penang. Similar to JobStreet, LinkedIn has its own industry classification, which does not fully match our priority industry branches. Yet, using a LinkedIn sales tool profiles have been mined within our scope of interest including higher-level qualifications and job function, position level, current location, industry type and seniority.

More than 20,000 profiles have been mined of persons who met the selection criteria. Of these, only a small percentage (15%) contained a verified email address. One important caveat to note is that LinkedIn does not trace the latest status of a LinkedIn account. A LinkedIn profile can be changed in real time where the employee profile registered in the database is assumed to be the most updated at the time of mining. Any revision made after the time of mining would not be registered in our database. The profiles collected resulted in a raw database in excel format, rows constituting work histories and other characteristics of individuals.

After some time, a number of significant drawbacks of the database became clear. Foremost, these involved the use by LinkedIn for classifications and categorisations of important variables that were inconsistent with either Malaysian classifications and/or categorisations that were deemed most useful for the study. Also, some necessary variables would have to be inserted manually, derived for other information in the profile. Cleaning and reworking the entire database will be tedious and time consuming, and we will not be able to complete this within the time frame given for this study.

It was therefore decided to consider only part of the database for processing and analysis. We decided to focus on only one destination industry, which is global business services (GBS) – or shared services. This was based on several reasons. First, it is a rather 'young' industry in Penang that is growing fast and that the state government is seeking to develop further. Second, part of this industry offers highly skilled jobs in a market that is already tight. At the same time,

sought-after in the market. Third, chain mobility effects in the labour market can be expected of growth of such a new industry; it will be interesting to consider the implications. After proper classification of industry of current job[29] to fit our own classification, employees currently in GBS can be extracted and further sampled. The final sample comprises some 300 individuals.

b) Questionnaire

A brief employee questionnaire has been developed consisting of several sections. The first section specifies the respondents' current employment status. The second section comprises questions pertaining to respondents' previous employment characteristics, including the commencement and end of each job, as well as its locations. It covers the respondents' last three jobs. This is followed by a third section with questions on drivers of job change and job satisfaction. Finally, personal particulars of the respondent covers the last section. A copy of the questionnaire is presented in Technical Report: Annex 13.

A digital form of this questionnaire has been made using Google Form. With a view to response, care was taken to ensure that completing the questionnaire would not take more than 20 minutes. For practical reasons, cases were selected from the sample that contained a verified email address. An invitation to fill the online questionnaire was sent to these persons by email. Of the approximately 100 persons invited, about half responded positively and returned the questionnaire.

c) Data analysis and presentation

This concerns the work histories and other characteristics of employees in the Global Business Services (GBS) sample of profiles and information obtained through the returned questionnaires.

3.5 Templates for discussion of (industry) results

To properly structure the presentation and discussion of our study findings, templates have been developed for Chapters 4–9 of this report. The templates for the macro-analysis, for individual industries within the scope of the study, and for mobility analysis are included in the Technical Report: Annex 11. We have opted to include recommendations following the findings into each of these templates, covering a skills strategy on issues raised in each perspective, which are macro-, meso- and micro-levels. Recommendations – general part and industry-specific parts – follow from analysis of issues, nature, and causes of skill shortages and gaps. We also highlight the roles of institutional framework.

3.6 Implementation obstacles and methodological limitations

This report is prepared based on information obtained from secondary research and information supplied to us by key stakeholders via interviews, surveys (questionnaires) and other means as highlighted in this chapter. Due to obstacles met during implementation, there are some limitations in regard to the findings presented in this study. The major limitations are as follows.

Component 1

The macro-level (statistical) data, obtained from secondary sources, should be handled with caution since there are several possible shortcomings. As it turns out, data are often incomplete, unavailable and/or inaccessible, which limit the possibility of analysing. Sometimes data are only available on national level and cannot be disaggregated to state-level. Besides that, multiple measures from different data sets were used, which can lead to different conclusions. Therefore, secondary data should be used and interpreted carefully (Veneri, 1999). Furthermore, it is important to note what some measures can and cannot explain. For example, when employers advertise a lot of vacancies, it does not necessarily mean they experience shortages (the company might be growing). Trends in vacancy data should therefore be evaluated together with other labour market indicators to understand the labour market

[29] GBS being a value chain activity in larger firms, in a number of cases GBS activities had been added to the portfolio of activities in already existing establishments in Penang, in the form of a new department (e.g. Intel's Shared Services Centre). This was taken into account in classification and labelling of current industry and the extraction/sampling procedures.

conditions and dynamics. Lastly, the vacancy data is a new way of obtaining information on the demand side of the labour market. Although JobStreet is the biggest job portal in Malaysia, it cannot be confirmed that it is representative of the entire labour market. Therefore, findings should be interpreted with caution and cannot be generalised. Besides that, this medium has not been used before, so procedures of filtering and analysing data are totally new. No reference projects or examples could be used in the approach, therefore the full process and steps taken in the analysis procedure are rather subjective. We have made sure to carefully explain each step in the analysis so that transparency is assured. In an environment where evidence for improving procedures is held in high esteem, drawing broad inferences from this observation merits careful interpretation.

Component 2

It has been noted that the validity, comprehensiveness, and information included in existing company databases in the state leave much to be desired. This should be given attention by authorities as updated and good quality databases are prerequisites for good research. Contact details of relevant firms derived from available company lists were often outdated, rendering the collection of current contact details of relevant firms difficult. Also, it was hard to obtain valid email addresses of human resource managers and directors. General email addresses of firms were of less value because they often get ignored. As a consequence, a thorough review of contact details has to been done by telephone or internet. This review was only partly successful. Many valid email addresses of HR managers were gathered through networking.

The employer survey used in this study to map out the skill situation of firms within industries also had its challenges. In the above, we have already discussed the limited response rate to the employer survey in the original and online formats. This can be ascribed to several reasons that negatively influenced the willingness of firms to participate. Various firms declined to participate in our study because they had participated in similar studies without much success. Next to that, they might assume that these types of studies are politically coloured to a certain extent. Also, employers exercise restraint in revealing recruitment strategies since they are all competing for the same pool of talented labour. No organisation wants to be open about the issues at hand. They indicate confidentiality of information and sensitivity of the topic as obstacles in participation. Furthermore, firms are still not entirely used to online surveys. Finally, questionnaires may have been considered lengthy and tedious to fill out. In this respect, the dilemma of the research team was the proper trade-off between information needed and time consumption of respondents. While survey is time consuming, oral interviews appear to be the best option. However, it remains difficult to set appointments with relevant firms as they are reserved in their availability and responses.

In the above, we have already explained how for the same reasons, participation in focus group discussions was rather meagre. This has implications for the depth of the qualitative information that could be collected and used.

Component 3

One major obstacle in labour mobility study is the difficulty of reaching employees (and the correct group). Companies often do not approve of their employees being approached, while employees do not want to be inconvenienced after work hours. We have already explained the issues surrounding the use of LinkedIn profiles to study mobility flows. These issues concern foremost classification/categorisation inconsistencies and incompleteness in regard to variables. Both issues are tedious and time consuming to resolve, especially if an initial database consists of thousands of cases. For all these reasons, a more limited set-up of the study of labour mobility gradually evolves. This is not necessarily a drawback as the component of mobility and skill-relatedness are meant to be exploratory and experimental, given the relative absence of labour mobility as a field of research in Malaysia. However, the set-up precludes a comprehensive picture of mobility patterns and skill relatedness, and what these do in the labour market. It is hoped that at some point,

a faster way of cleaning and reworking databases derived from LinkedIn becomes possible as LinkedIn starts to require members listing information in a way that it corresponds to international (classification and statistical) standards.

Although good quality and interesting software has become available to graphically depict mobility flows in novel ways, this software (Circos has been used for this purpose) is highly complex and difficult to operate.