

Profiling Thai Labor Supply Through a Million Online Job Seekers

บทคัดย่อ

งานวิจัยนี้ใช้ข้อมูลประวัติการศึกษาและประวัติการทำงานในระดับบุคคล จากข้อมูลประวัติการทำงานออนไลน์ (Online resume) ของเว็บไซต์ผู้ให้บริการจัดหางานสองรายในประเทศไทย ข้อมูลรวมสำหรับใช้ในการวิเคราะห์ ทั้งหมด 1,324,225 คน โดยแบ่งออกเป็นชุดข้อมูลที่มีเฉพาะผู้ที่มีประสบการณ์ทำงาน ซึ่งสามารถนำไปจัดโครงสร้างข้อมูลให้เป็นแบบ unbalanced panel data แบบรายเดือน ได้เป็นจำนวนเท่ากับ 710,788 คน และมีจำนวนข้อมูลทั้งสิ้น 35,924,325 คน-เดือน คุณลักษณะต่างๆของข้อมูล ORE ซึ่งให้เห็นถึงข้อควรระวังในการประยุกต์ใช้ข้อมูล เนื่องจากข้อมูลส่วนใหญ่เป็นข้อมูลที่มาจากกลุ่มที่มีอายุน้อย (20-29 ปี) และมีการศึกษาที่สูงกว่าประชากรโดยทั่วไป อย่างไรก็ตาม ด้วยจำนวนของกลุ่มตัวอย่างที่มีขนาดใหญ่ ก็เป็นโอกาสให้นักวิจัยสามารถใช้ประโยชน์จากข้อมูลที่มีความครอบคลุมทั้งในเชิงพื้นที่และช่วงเวลา ตลอดจนรายละเอียดข้อมูลการศึกษาและประวัติการทำงาน เพื่อเพิ่มประสิทธิภาพในการออกแบบการทดลองกับข้อมูล

ข้อมูลซึ่งให้เห็นว่ารูปแบบการเลือกสาขาวิชาในการศึกษาระดับปริญญาตรี รวมถึงกลุ่มสาขาอาชีพของแรงงานไทย นั้น ไม่มีการเปลี่ยนแปลงนักในช่วง 40 ปีที่ผ่านมา โดยส่วนใหญ่ก็ยังนิยมที่จะเลือกทำการศึกษากลุ่มสาขาวิชาสังคมศาสตร์ และบริหารธุรกิจ และสาขาวิชาวิศวกรรมศาสตร์ โดยที่กลุ่มวิชาและกลุ่มสาขาอาชีพที่เกี่ยวข้องกับเทคโนโลยีก็เริ่มได้รับความสนใจมากขึ้นอย่างต่อเนื่อง นอกจากนี้ คนรุ่นใหม่มีแนวโน้มที่จะไม่ทำงานในองค์กรใดองค์กรหนึ่งนานๆ และคนในรุ่นก่อนหน้าก็มีแนวโน้มที่จะปรับตัวตามสถานการณ์ที่เปลี่ยนแปลงไปในตลาดแรงงานเช่นกัน

1. ที่มา

ในปี 2562 ประเทศไทยมีจำนวนผู้ใช้งานอินเทอร์เน็ตทั้งสิ้น 50,080,000 ราย ซึ่งเป็นการเพิ่มขึ้นกว่า 339 เปอร์เซ็นต์ เมื่อเทียบกับจำนวนผู้ใช้งานอินเทอร์เน็ตในปี 2550¹ การเพิ่มขึ้นอย่างรวดเร็วของผู้ใช้งานอินเทอร์เน็ตในประเทศไทยในช่วงทศวรรษที่ผ่านมา ส่งผลให้เกิดความเปลี่ยนแปลงในหลายภาคส่วนในระบบเศรษฐกิจ ข้อมูลขนาดใหญ่บนเว็บไซต์หรือแอปพลิเคชันซึ่งบันทึกข้อมูลกิจกรรม พฤติกรรม ตลอดจนช่วงเวลาในการใช้บริการของผู้เข้ามาใช้บริการ ต่างถูกนำมาวิเคราะห์และใช้ประโยชน์ในงานวิจัยด้านสังคมศาสตร์เพิ่มมากขึ้นเรื่อยๆ (Askitas & Zimmermann, 2015) หนึ่งใน การเปลี่ยนแปลงที่สำคัญ โดยเฉพาะอย่างยิ่งกับตลาดแรงงาน คือ การให้บริการจัดหางานทางอินเทอร์เน็ต (Internet Job Portal: IJP) และ พฤติกรรมการหางานทางอินเทอร์เน็ต² ซึ่งได้กลายมาเป็นกลไกที่สำคัญของตลาดแรงงานในปัจจุบัน (Freeman, 2002) นอกจากนี้ ฐานข้อมูลของผู้ให้บริการจัดหางานทางอินเทอร์เน็ตแต่ละราย ซึ่งบันทึกข้อมูลของผู้เข้ามาใช้บริการ ไม่ว่าจะเป็นผู้ที่กำลังหางาน (Unemployed Searchers) ผู้ที่กำลังต้องการเปลี่ยนงาน (On-the-job Searchers) หรือแม้แต่องค์กรธุรกิจที่กำลังต้องการแรงงาน ต่างก็ถูกนำมาวิเคราะห์และใช้ประโยชน์ในงานวิจัยเพิ่มมากขึ้นเรื่อยๆ

จุดประสงค์ของงานวิจัยนี้ประกอบด้วยสองส่วน คือ (1) เพื่อนำเสนอข้อมูลด้านอุปทานของตลาดแรงงานที่เกิดขึ้นบนอินเทอร์เน็ต ที่ในปัจจุบันได้กลายมาเป็นหนึ่งในสื่อกลางหลักสำหรับการหางาน (อุปทาน) และหาแรงงาน (อุปสงค์) ของประเทศไทย และ (2) เพื่อนำเสนอชุดข้อมูลที่เป็นทางเลือกสำหรับงานวิจัยในประเทศไทยต่อไปในอนาคต ซึ่งงานวิจัยนี้ได้สร้างจากข้อมูลประวัติออนไลน์ (Online Resume: ORE) ที่ครอบคลุมทั้งข้อมูลประวัติการศึกษาและข้อมูลประวัติการทำงานในอดีตในระดับบุคคล ซึ่งเกิดจากการที่แรงงาน (ในฐานะผู้ใช้บริการ) ได้เข้าไปใช้บริการที่มักถูกเรียกว่า “บริการรับฝากข้อมูลประวัติ” (Resume Posting Service) ซึ่งเป็นการอำนวยความสะดวกให้กับทั้งผู้ใช้บริการที่เป็นแรงงานที่กำลังมองหา งาน (โดยสามารถฝาก ORE ไว้กับเว็บไซต์หรือแอปพลิเคชันของ IJP แต่ละราย) และผู้ใช้บริการที่เป็นองค์กรธุรกิจ (โดยสามารถเข้ามาทำการคัดเลือกผู้ที่มีประวัติที่น่าสนใจ และติดต่อเพื่อทำการสัมภาษณ์ต่อไป)

ข้อมูลในงานวิจัยนี้นำมาใช้แตกต่างจากชุดข้อมูลที่นิยมนำมาใช้ในงานวิจัยที่ผ่านมา ซึ่งมักจะใช้ข้อมูลการเข้าไปค้นหาในบอร์ดรับสมัครงาน (Mamertino & Sinclair, 2016) ข้อมูลการยื่นสมัครงานออนไลน์ (Shen & Kuhn, 2013; Brown & Matsa, 2016, 2019; Marinescu, 2017; Marinescu & Rathelot, 2018) และข้อมูลการประกาศรับสมัครงาน (Kuhn & Shen, 2013; Marinescu, 2017; Deming & Kahn, 2018; Hershbein & Kahn, 2018; Marinescu & Rathelot, 2018; Turrell, Speigner, Djumalieva, Copple, & Thurgood, 2019)

ในงานวิจัยที่ผ่านมา มีเพียง Kurekova & Zilincikova (2018) ที่ใช้ข้อมูล ORE ของแรงงานในประเทศสโลวาเกีย เพื่อมาศึกษาผลกระทบที่เกิดขึ้นต่อตลาดแรงงานในประเทศ เมื่อแรงงานชาวสโลวาเกียที่มีประสบการณ์ทำงานในต่างประเทศ กลับสู่บ้านเกิด และงานวิจัยซึ่งศึกษาตลาดแรงงานในประเทศไทยของ Lekfuangfu, Nakavachara, & Sawaengsuksant (2017) เท่านั้น ที่เคยมีการใช้ประโยชน์จากข้อมูล ORE เช่นเดียวกันกับงานวิจัยนี้ อย่างไรก็ตาม สิ่งที่เพิ่มเติมมาสำหรับงานวิจัยนี้ ก็คือการนำเสนอแนวทางการจัดการข้อมูลตัวอักษร (text data) เพื่อนำข้อมูลตัวอักษรที่ได้มาใช้ประโยชน์ และบวกกับธรรมชาติของข้อมูล ORE ที่มีลักษณะเป็นทะเบียนประวัติหรือบันทึกประวัติ ซึ่งมีทั้งข้อมูลส่วนตัว ข้อมูลการศึกษา

¹ ข้อมูลจำนวนผู้ใช้งานอินเทอร์เน็ตรวมแบบประจำที่และแบบเคลื่อนที่ จากเว็บไซต์สำนักงานคณะกรรมการกิจการกระจายเสียง กิจการโทรทัศน์ และกิจการโทรคมนาคมแห่งชาติ

² คำว่า “การหางานทางอินเทอร์เน็ต” ในงานวิจัยนี้มีความหมายครอบคลุมถึง กิจกรรมการหางานของแรงงาน และการหาแรงงานขององค์กรธุรกิจ ไม่ว่าจะกิจกรรมเหล่านี้จะเกิดขึ้นบนเว็บไซต์หรือแอปพลิเคชันสำหรับโทรศัพท์เคลื่อนที่ซึ่งพัฒนาโดยผู้ให้บริการจัดหางานแต่ละราย

และ ข้อมูลประวัติการทำงานของกลุ่มตัวอย่าง ซึ่งทำให้งานวิจัยนี้เป็นงานวิจัยแรกๆ ที่สามารถประยุกต์ใช้แนวทางการจัดการข้อมูลตัวอักษรเพื่อให้ได้มาซึ่งข้อมูลที่อยู่ในรูปแบบของ panel data ซึ่งน่าจะเป็นประโยชน์ต่อการรวบรวมและจัดโครงสร้างข้อมูลในอนาคต เนื่องจากในปัจจุบันนี้ ข้อมูลตัวอักษรถือได้ว่าเป็นประเภทข้อมูลที่เกิดขึ้นเป็นจำนวนมากและยังเกิดขึ้นตลอดเวลา (เช่น ข้อความจากการโพสต์เฟสบุ๊ก การรีวิวกความพึงพอใจในเว็บไซต์ การโพสต์กระทู้ เป็นต้น) ซึ่งข้อมูลเหล่านี้ยังไม่ค่อยได้มีการนำมาใช้ประโยชน์มากนัก โดยเฉพาะอย่างยิ่งกับงานวิจัยทางด้านเศรษฐศาสตร์

องค์ประกอบของบทความมีดังต่อไปนี้ ส่วนที่ 2 จะอธิบายรายละเอียด แหล่งที่มาของข้อมูลประวัติการทำงานออนไลน์ ขั้นตอนการคัดกรอง และขั้นตอนการจัดโครงสร้างข้อมูล ตลอดจนคุณลักษณะที่สนใจของข้อมูล ORE ส่วนที่ 3 นำเสนอข้อมูลคุณลักษณะที่โดดเด่น (stylized facts) ของอุปทานแรงงานไทยที่เกิดบนสื่อกลางประเภทอินเทอร์เน็ต ส่วนต่อมา ทำการอภิปรายถึงปัญหาในการใช้ข้อมูลจากการสำรวจภาวะการทำงานของประชากร (Labor Force Survey: LFS) เพื่อมาเป็นตัวแทนประชากรของข้อมูล รวมถึง ข้อควรระวังในการใช้ข้อมูล ORE และ ส่วนที่ 5 นำเสนอแนวทางการใช้ประโยชน์ข้อมูล ORE และส่วนสุดท้าย ทำการอภิปรายข้อสรุปและแนวทางในการพัฒนาข้อมูล ORE เพิ่มเติมในอนาคต

2. ข้อมูล ORE: ที่มา การคัดกรองและการจัดโครงสร้างข้อมูล และการเป็นตัวแทนประชากร

2.1 ที่มาของข้อมูล

งานวิจัยนี้สร้างชุดข้อมูลประวัติการศึกษาและประวัติการทำงานในอดีต ในระดับบุคคล โดยใช้ฐานข้อมูล ORE ซึ่งถูกตั้งค่าให้สามารถเข้าถึงได้แบบสาธารณะ ของเว็บไซต์ผู้ให้บริการจัดหางานจำนวนสองรายในประเทศไทย (ซึ่งต่อไปนี้จะเรียกว่า IJP-A และ IJP-B) งานวิจัยนี้สร้าง web scraper เพื่อทำการดึงข้อมูลจากเว็บไซต์โดยใช้ R Programming Language (R Core Team, 2017) โดยทำการดึงข้อมูลในช่วงเดือนเมษายน 2563 จากข้อมูลประวัติการทำงานทั้งหมดในฐานข้อมูลของ IJP-A และ IJP-B รวมทั้งสิ้น 2,589,904 User IDs (UID) (จำนวนประวัติการทำงานจาก IJP-A และ IJP-B เท่ากับ 1,935,732 และ 654,172 UID ตามลำดับ)

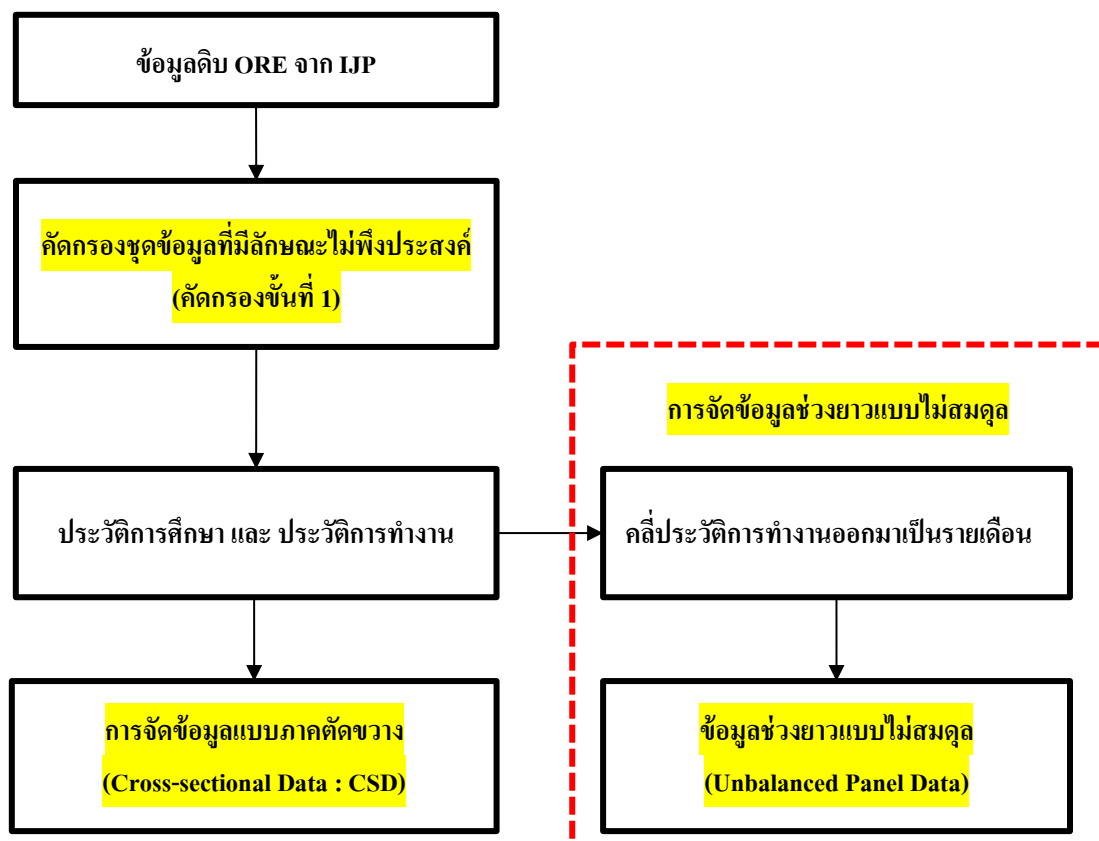
ข้อมูล ORE จากฐานข้อมูลของเว็บไซต์ผู้ให้บริการทั้งสอง มีลักษณะที่คล้ายคลึงกัน และสามารถแบ่งออกได้เป็นสามส่วน คือ ข้อมูลประวัติส่วนตัว ข้อมูลประวัติการศึกษา และข้อมูลประสบการณ์ทำงาน³ โดยข้อมูลประวัติส่วนตัวประกอบด้วย วันเดือนปีที่มีการแก้ไขประวัติครั้งล่าสุด จังหวัดที่อยู่ปัจจุบัน วันเดือนปีเกิด ส่วนสูง อายุ การเกณฑ์ทหาร งาน และประเภทของงานที่ต้องการ ระดับเงินเดือนที่คาดหวัง โดยในส่วนของข้อมูลประวัติศึกษานั้น เว็บไซต์ผู้ให้บริการจัดหางานทั้งสองราย จะอนุญาตให้ผู้ใช้บริการใส่รายละเอียดประวัติการศึกษาสูงสุดสองอันดับแรกเช่นเดียวกัน โดยรายละเอียดประวัติการศึกษาจะประกอบไปด้วย ระดับการศึกษา ชื่อสถานศึกษา ชื่อวุฒิการศึกษา สาขาวิชา และ ข้อมูลระดับคะแนนเฉลี่ย (Grade Point Average: GPA) อย่างไรก็ตาม เว็บไซต์ผู้ให้บริการทั้งสองรายจะอนุญาตให้ผู้ใช้บริการแจ้งข้อมูลปีที่จบการศึกษาเฉพาะการศึกษาสูงสุดเท่านั้น สำหรับข้อมูลประวัติการทำงานนั้น จะประกอบไปด้วย รายละเอียดเดือนและปีที่เริ่มทำงาน เดือนและปีสุดท้ายที่ทำงาน ชื่อองค์กรที่เคยทำงาน ที่อยู่ขององค์กรที่เคยทำงาน ตำแหน่งสุดท้ายก่อนออกจากงาน เงินเดือนเดือนสุดท้ายที่ได้รับ และ รายละเอียดลักษณะงานที่ทำ อย่างไรก็ตาม ความแตกต่างที่สำคัญในส่วน

³ เพื่อเป็นการประหยัดพื้นที่บทความ ข้อมูลเพื่อประกอบการพิจารณาอื่นๆ เช่น ข้อมูลประวัติการอบรม การฝึกงาน และเกียรติประวัติที่ผ่านมา ที่มีปรากฏในข้อมูล ORE แต่ไม่ได้ใช้ในงานวิจัยนี้ จึงไม่ได้ถูกอภิปรายถึง

ประวัติการทำงาน ก็คือ จำนวนข้อมูลประวัติการทำงานในอดีต โดยที่ IJP-A และ IJP-B อนุญาตให้ผู้ใช้บริการสามารถให้ข้อมูลประวัติการทำงานในอดีตได้สูงสุดสามแห่ง และ สองแห่ง ตามลำดับ

2.2 การคัดกรองและการจัดโครงสร้างข้อมูล

ข้อมูลประวัติการทำงานเป็นประเภทข้อมูลที่เป็นตัวอักษร (Text data) และอยู่ในรูปแบบของข้อมูลที่ยังไม่ได้ถูกจัดโครงสร้าง (Unstructured data) ดังนั้น ก่อนที่จะนำข้อมูลไปวิเคราะห์ในขั้นต่อไป เราจึงจำเป็นต้องนำข้อมูลดิบที่ได้มาเหล่านั้น เข้าสู่กระบวนการทำความสะอาดข้อมูล (Data cleaning) และกระบวนการจัดโครงสร้าง (Data manipulation) อย่างไรก็ตาม เนื่องจากข้อมูล ORE มีโครงสร้างและลำดับของรายละเอียดที่แน่นอนและเป็นไปตามแบบฟอร์มที่ผู้ให้บริการได้จัดเตรียมไว้ งานวิจัยนี้จึงประยุกต์ใช้เทคนิค Regular Expression (Regex) ซึ่งเป็นวิธีการที่ใช้ในศาสตร์ทางวิทยาการคอมพิวเตอร์ (Computer Science) และการประมวลผลภาษาธรรมชาติ (Natural Language Processing) ในการกำหนดรูปแบบของคำหรือกลุ่มคำ เพื่อใช้ในการค้นหาหรือข้อความที่ต้องการในการจัดการกับโครงสร้างของข้อมูลที่เป็นประเภทข้อมูลตัวอักษรเหล่านั้น รูปที่ 1 แสดงแผนภาพสรุปขั้นตอนการคัดกรองขั้นที่ 1 และการจัดโครงสร้างข้อมูล



รูปที่ 1: แผนภาพสรุปขั้นตอนการคัดกรองขั้นที่ 1 และการจัดโครงสร้างข้อมูล

เนื่องจากข้อมูล ORE นั้น เกิดจากการที่ผู้ใช้บริการทำการพิมพ์ข้อมูลเข้าไปในแบบฟอร์มออนไลน์ตามที่เว็บไซต์ผู้ให้บริการได้จัดเตรียมไว้ด้วยตนเอง ด้วยเหตุนี้ จึงมีโอกาสสูงที่จะเกิดข้อผิดพลาดในขั้นตอนการใส่ข้อมูลได้ ด้วยเหตุนี้

หลังจากขั้นตอนการดึงข้อความในส่วนที่ต้องการด้วยเทคนิค RegEx แล้ว เราจึงต้องทำการตรวจสอบความถูกต้องของข้อมูลอีกครั้งหนึ่ง ในการคัดกรองขั้นที่หนึ่งนี้ เราจะทำการค้นหาชุดข้อมูลที่มีลักษณะไม่พึงประสงค์ เช่น

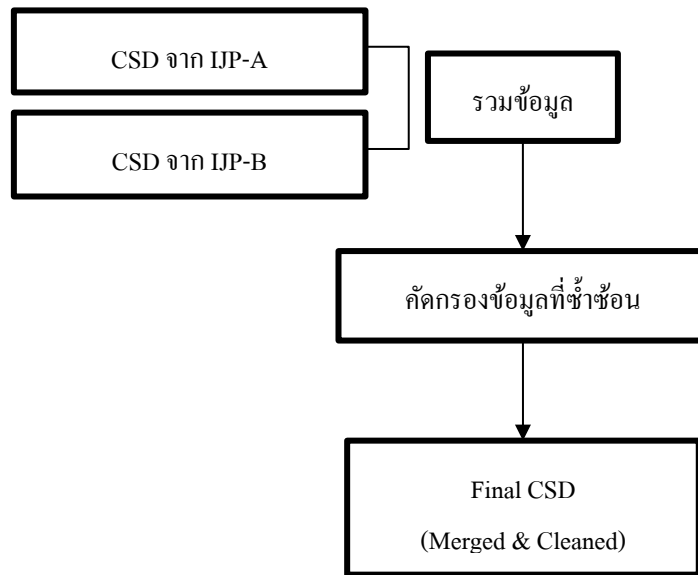
- ชุดข้อมูลที่ใช้บริการไม่ได้ให้ข้อมูลประวัติการศึกษา
- ชุดข้อมูลที่ใช้บริการที่กำลังศึกษาอยู่
- ชุดข้อมูลที่มีรายละเอียดที่ขัดแย้งกันเอง ยกตัวอย่างเช่น ระดับการศึกษาสูงสุด คือปริญญาโท แต่ระบุพฤติกรรมการศึกษาว่าเป็น วิทยาศาสตร์บัณฑิต ซึ่งเป็นวุฒิการศึกษาของระดับปริญญาตรี ทั้งนี้ การคัดกรองยังได้ได้อาศัยข้อมูลชื่อปริญญาตามประกาศกระทรวงศึกษาธิการ เรื่อง หลักเกณฑ์การกำหนดชื่อปริญญา พ.ศ. 2559 ซึ่งกำหนดชื่อปริญญาของทุกสาขาวิชาของทุกระดับการศึกษาในระดับอุดมศึกษาของประเทศไทย มาช่วยในการคัดกรองความถูกต้องของข้อมูล
- ชุดข้อมูลที่มีการบ่งชี้ถึงการปลอมแปลงหรือน่าสงสัย ยกตัวอย่างเช่น ระดับการศึกษาสูงสุด คือปริญญาโท แต่ระบุชื่อสถานศึกษาว่าเป็น “aaaaaa” หรือการระบุระดับคะแนนเฉลี่ย เท่ากับ 5.5
- ชุดข้อมูลที่มีรายละเอียดขัดแย้งกับกฎระเบียบหรือหลักเกณฑ์ ซึ่งถูกประกาศโดยหน่วยงานที่รับผิดชอบ เช่น เกณฑ์มาตรฐานหลักสูตรระดับปริญญาตรี เกณฑ์มาตรฐานคุณวุฒิอาชีวศึกษา ทั้งระดับประกาศนียบัตรวิชาชีพ (ปวช.) และวิชาชีพชั้นสูง (ปวส.) ยกตัวอย่างเช่น ระดับการศึกษาสูงสุด คือ ปวส. แต่ระบุระดับคะแนนเฉลี่ยเท่ากับ 1.8 ซึ่งขัดกับเกณฑ์มาตรฐานคุณวุฒิอาชีวศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ซึ่งกำหนดให้ระดับคะแนนเฉลี่ยเพื่อสำเร็จการศึกษาตามหลักสูตร จะต้องไม่ต่ำกว่า 2.0

อย่างไรก็ตาม ในส่วนของข้อมูลประวัติการทำงานในอดีตนั้น การคัดกรองชุดข้อมูลที่มีลักษณะไม่พึงประสงค์ออกไป จะเป็นการคัดกรองที่มุ่งเน้นเฉพาะในส่วนของประวัติการทำงานที่ไม่ได้ระบุรายละเอียดช่วงระยะเวลาที่ทำงาน (เดือนและปีที่เริ่มทำงาน เดือนและปีสุดท้ายที่ทำงาน)⁴ ประวัติการทำงานที่เป็นส่วนหนึ่งของหลักสูตรการศึกษา (การฝึกงานหรือสหกิจศึกษา) ประวัติการทำงานที่ไม่ระบุตำแหน่งการทำงานที่ชัดเจน และประวัติการทำงานที่ไม่ระบุรายละเอียดใดๆโดยสิ้นเชิง ทั้งนี้ ชุดข้อมูลที่มีความผิดปกติต่างๆที่ถูกรวบรวมในขั้นนี้ จะยังไม่ถูกลบหรือคัดกรองออกไปทันที แต่จะถูกบันทึกไว้ในรูปแบบของตัวแปรที่มีชื่อแตกต่างกันไป ตามลักษณะของความผิดปกติแบบต่างๆ ที่ได้ถูกโปรแกรมให้ค้นหาโดยอัตโนมัติ เพื่อประโยชน์ในการตรวจสอบในภายหลัง และความสะดวกต่อการวิเคราะห์ในประเด็นที่ไม่มีความจำเป็นจะต้องใช้ประวัติการทำงานอดีต

ข้อมูลที่ถูกคัดกรองจากขั้นที่หนึ่งเรียบร้อยแล้ว จะถูกนำไปจัดเรียงโครงสร้างให้อยู่ในรูปแบบของตารางเพื่อสะดวกต่อการวิเคราะห์ ทั้งนี้ เพื่อเป็นการใช้ประโยชน์จากข้อมูลให้ได้สูงสุด งานวิจัยนี้ได้จัดทำโครงสร้างของข้อมูลประวัติการทำงานออกมาสองประเภท คือ ข้อมูลแบบภาคตัดขวาง (cross-sectional data) ซึ่งจัดเก็บข้อมูลรายละเอียดประวัติการศึกษาและประสบการณ์ในการทำงานรายบุคคล และ ข้อมูลช่วงยาวแบบไม่สมดุล (unbalanced panel data) ซึ่งใช้ชุดข้อมูลเดียวกันกับชุดข้อมูล cross-section แต่มีการจัดโครงสร้างรายละเอียดของประสบการณ์ในการทำงานให้เป็นแบบราย

⁴ บ่อยครั้งที่เรามักจะพบว่า ผู้ใช้บริการทำการระบุเดือนและปีสุดท้ายในการทำงานในองค์กรล่าสุดว่า “ปัจจุบัน” ซึ่งในกรณีเช่นนี้ งานวิจัยนี้จะใช้ข้อมูลวันเดือนปีที่มีการแก้ไขประวัติครั้งล่าสุด มาเป็นตัวแทน (Proxy) ในการระบุช่วงเวลาเดือนและปีสุดท้ายในการทำงานในองค์กรล่าสุด อย่างไรก็ตาม หากพบว่ามีกรณีเดือนและปีสุดท้ายว่า “ปัจจุบัน” มากกว่าหนึ่งองค์กรในประวัติการทำงานในอดีต ประสบการณ์ทำงานในองค์กรเหล่านั้น จะถูกคัดกรองออกไป

เดือน โดยใช้ประโยชน์จากข้อมูลช่วงเวลาที่เคยทำงานในแต่ละองค์กรในอดีต ทำให้เราได้มาซึ่งชุดข้อมูลประวัติการทำงานที่มีความถี่สูง แต่จำนวนข้อมูลของแต่ละตัวอย่างมีความแตกต่างกันไป ขึ้นอยู่กับประสบการณ์ในการทำงานของแต่ละคน



รูปที่ 2: แผนภาพขั้นตอนการคัดกรองขั้นที่ 2

รูปที่ 2 แสดงแผนภาพสรุปขั้นตอนการคัดกรองขั้นต่อมา หลังจากการจัดโครงสร้างข้อมูล โดยในการคัดกรองขั้นที่ 2 นั้น เนื่องจากบริการฝากประวัติการทำงานเป็นบริการที่ไม่ต้องเสียค่าใช้จ่าย จึงมีความเป็นไปได้ที่จะมีการฝากประวัติการทำงานที่ซ้ำซ้อน ไม่ว่าจะเป็นการเปิดบัญชีผู้ใช้งานมากกว่าหนึ่งบัญชีในเว็บไซต์ผู้ให้บริการรายเดียวหรือการฝากประวัติการทำงานกับเว็บไซต์ที่ให้บริการจัดหางานมากกว่าหนึ่งแห่ง ด้วยเหตุนี้ หลังจากการจัดโครงสร้างของข้อมูลแล้ว เราจำเป็นต้องทำการคัดกรองข้อมูลที่ซ้ำซ้อนเหล่านี้ออกไป โดยงานวิจัยนี้ได้ใช้ข้อมูลวันเดือนปีเกิด เพศ ชื่อสถานศึกษา⁵ และระดับคะแนนเฉลี่ย มาเป็นตัวบ่งชี้ความซ้ำซ้อนของข้อมูลของผู้ใช้บริการแต่ละราย ซึ่งทำให้จะต้องมีการตรวจสอบความน่าเชื่อถือของข้อมูลวันเดือนปีเกิดของผู้ใช้ โดยใช้การคำนวณอายุ ณ วันที่มีการแก้ไขประวัติครั้งล่าสุดและอายุ ณ ปีที่จบการศึกษาของระดับการศึกษาสูงสุดของผู้ใช้บริการ และนำมาเปรียบเทียบกับอายุน้อยที่สุดที่เป็นไปได้ที่จะจบการศึกษาในระดับที่ได้ระบุไว้⁶ ซึ่งเป็นเหตุให้จำนวนข้อมูล ORE ของ IJP-A และ IJP-B ภายหลังขั้นตอนการคัดกรองประวัติการทำงานที่ซ้ำซ้อนเท่ากับ 1,079,166 และ 427,209 UID ตามลำดับ โดยหลังจากการผนวกรวมชุดข้อมูลจากเว็บไซต์ผู้ให้บริการทั้งสอง และทำการคัดกรองข้อมูลที่มีความเป็นไปได้ที่จะมาจากผู้ที่กำลังหางานคนเดียวกันที่มีอยู่ในข้อมูล ORE ของทั้ง IJP-A และ IJP-B ออก ทำให้เหลือจำนวนข้อมูลรวมสำหรับการวิเคราะห์ ทั้งหมด 1,324,225 คน โดยในจำนวนนี้ แบ่งออกเป็นชุดข้อมูลที่สามารนำไปจัดโครงสร้างข้อมูลให้เป็นแบบ unbalanced panel data แบบรายเดือน ได้เป็นจำนวน

⁵ ข้อมูลชื่อสถานศึกษามากกว่า 239,000 ชื่อ ซึ่งหลายชื่อเป็นสื่อความหมายเดียวกันแต่มีการสะกดที่ผิดพลาด เช่น “โรงเรียนวัดสุทธิ” “มหาลัย เกษศาสตร์” ทำให้ข้อมูลชื่อสถานศึกษาเหล่านี้ต้องถูกแก้ไขก่อนที่จะนำไปค้นหาตำแหน่งที่ตั้ง (ละติจูด/ลองจิจูด) ซึ่งทั้งสองขั้นตอนนี้ ผู้วิจัยได้ใช้ ความสามารถของ Places API และ Geocoding API ซึ่งพัฒนาโดย Google ซึ่งสามารถทำความเข้าใจโครงสร้างของคำรวมถึงเจตนาของคำค้นหา และทำการพยากรณ์คำที่ผู้ใช้งานจะค้นหาได้อย่างแม่นยำ

⁶ งานวิจัยนี้ใช้อายุน้อยที่สุดที่เป็นไปได้สำหรับจบการศึกษาระดับประถมศึกษา มัธยมศึกษาตอนต้น มัธยมศึกษาตอนปลาย ประกาศนียบัตรวิชาชีพ ประกาศนียบัตรวิชาชีพชั้นสูง ปริญญาตรี ปริญญาโท และปริญญาเอก คือ 12 ปี 15 ปี 18 ปี 18 ปี 20 ปี 22 ปี 24 ปี และ 28 ปี ตามลำดับ

เท่ากับ 710,788 คน และมีจำนวนข้อมูลทั้งสิ้น 35,924,325 คน-เดือน (man-month)

ทั้งนี้ ปัญหาที่สำคัญที่พบในการใช้งานข้อมูล ORE ของงานวิจัยนี้ คือ ปัญหาการแบ่งประเภทอาชีพ เนื่องจากชื่อตำแหน่งที่ใช้ในแต่ละองค์กรธุรกิจมีความแตกต่างกันออกไป ถึงแม้ว่าลักษณะและรายละเอียดการทำงานจะคล้ายกัน เช่น “พนักงานขาย” “เซลล์ขาย” “Sales Executive” “Account Executive” “พนักงานส่งเสริมการขาย” “Sales Representative” หรือแม้แต่ “Sales Engineer” ซึ่งหน้าที่หลักของตำแหน่งเหล่านี้คือการขายสินค้าหรือบริการขององค์กรธุรกิจเหมือนกัน โดยถึงแม้ว่าในการวิเคราะห์นั้น เราจะสามารถใช้ข้อมูลประเภทหรือชนิดของงานซึ่งถูกเลือกโดยตัวผู้ให้บริการเองในขั้นตอนการฝากประวัติ เพื่อมาช่วยในการบ่งชี้ประเภทหรือชนิดของงานที่แรงงานกำลังหรือมีความต้องการที่จะทำได้ก็ตาม แต่ปัญหาที่พบต่อมาก็คือ การเลือกประเภทหรือชนิดของงานที่ไม่ตรงกับสายงานที่กำลังทำหรือเคยทำมาเลย และด้วยปริมาณเป็นจำนวนมากของข้อมูลตัวอักษรที่ใช้งานวิจัยนี้ จึงทำให้มีความจำเป็นที่จะต้องพัฒนาอัลกอริทึมเพิ่มขึ้นมาใหม่เพื่อเป็นเครื่องมือที่จะนำมาวิเคราะห์ความเป็นไปได้ในการจัดประเภทหรือชนิดของงานที่ระบุในข้อมูล ORE ต่อไป ด้วยเหตุนี้ งานวิจัยนี้จึงยังไม่สามารถนำข้อมูลในส่วนนี้มาใช้ได้

2.3 คุณลักษณะที่น่าสนใจ

ตารางที่ 1 แสดงรายละเอียดคุณลักษณะที่น่าสนใจของข้อมูล ORE เฉพาะกลุ่มตัวอย่างที่มีประสบการณ์ทำงาน ซึ่งกว่า 63% ของข้อมูล คือ กลุ่มตัวอย่างที่จบการศึกษาในระดับปริญญาตรี รองลงมาคือ ปวส. ม.6 และ ปวช. ซึ่งมีสัดส่วนเท่ากับ 17% 8% และ 6% ตามลำดับ โดยกว่า 66% ของกลุ่มตัวอย่างทุกระดับการศึกษายังมีสถานภาพโสด โดยเฉพาะอย่างยิ่งกลุ่มตัวอย่างในระดับปริญญาตรีซึ่งมีสัดส่วนของคนที่ยังไม่ได้แต่งงานกว่า 69% โดยสัดส่วนของกลุ่มตัวอย่างเพศหญิงรวมทุกระดับการศึกษามีกว่า 63% ซึ่งผู้จบการศึกษาสูงสุดในระดับปริญญาตรีมีสัดส่วนของกลุ่มตัวอย่างเพศหญิงสูงที่สุดถึง 68% ซึ่งใกล้เคียงกับสัดส่วนของนักศึกษาหญิงที่กำลังศึกษาในระดับปริญญาตรีชั้นปีที่ 4 ในช่วงทศวรรษที่ผ่านมาจากกระทรวงศึกษาธิการซึ่งมีสัดส่วนประมาณ 60%⁷ อย่างไรก็ตาม ผู้จบการศึกษาสูงสุดในระดับประถมศึกษา มีสัดส่วนของกลุ่มตัวอย่างเพศหญิงต่ำที่สุดเพียง 28% ซึ่งแตกต่างอย่างมากจากสัดส่วนของนักเรียนหญิงที่กำลังศึกษาในระดับประถมศึกษาปีที่หกในช่วงทศวรรษที่ผ่านมา ที่มีสัดส่วนเท่ากับ 48%

ค่าเฉลี่ยของอายุ ณ ปีที่จบการศึกษาของกลุ่มตัวอย่างที่มีระดับการศึกษาสูงสุดในระดับประถมศึกษา จนถึง ระดับปริญญาตรี ของข้อมูล ORE ที่ใกล้เคียงกับเกณฑ์อายุเฉลี่ยของผู้จบการศึกษาหากพิจารณาตามเกณฑ์การศึกษาภาคบังคับของประเทศไทย ที่กำหนดให้เด็กทุกคนต้องเข้าศึกษาในระดับประถมศึกษาชั้นปีที่ 1 เมื่ออายุย่างเข้าปีที่เจ็ด บ่งชี้ว่ากลุ่มตัวอย่างของข้อมูล ORE โดยเฉลี่ยแล้วน่าจะมีการศึกษาที่ต่อเนื่อง อย่างไรก็ตาม สำหรับการศึกษาในระดับบัณฑิตศึกษา (ปริญญาโท และปริญญาเอก) นั้น โดยเฉลี่ยแล้วกลุ่มตัวอย่างมีแนวโน้มที่จะไม่ได้ทำการศึกษาต่อเนื่องเมื่อพิจารณาจากระยะเวลาปกติของหลักสูตรเหล่านี้โดยทั่วไป ซึ่งเป็นลักษณะที่เห็นได้ทั่วไปในปัจจุบัน โดยเฉพาะในระดับปริญญาโท **ตารางที่ 1** ยังแสดงข้อมูลค่าเฉลี่ยอายุของผู้ให้บริการทางอินเทอร์เน็ตของข้อมูล ORE ในอีกสองช่วงเวลา คือ อายุเฉลี่ย ณ ปี พ.ศ. 2562 และ อายุเฉลี่ย ณ วันที่มีการแก้ไขข้อมูลล่าสุด ซึ่งผู้วิจัยทำการคำนวณขึ้นมาเพื่อชี้ให้เห็นถึงแนวโน้มความห่างของระยะเวลาในการปรับปรุง (update) ข้อมูลประวัติการทำงาน โดยในกลุ่มตัวอย่างรวมนั้น พบว่า กลุ่มตัวอย่างในข้อมูล ORE นี้ โดยเฉลี่ยแล้ว ไม่มีการแก้ไขข้อมูลประวัติการทำงานใดๆ ในช่วงสี่ปีที่ผ่านมา ซึ่งพฤติกรรมการแก้ไขข้อมูลดังกล่าว ก็มีความ

⁷ งานวิจัยนี้จึงใช้สัดส่วนของนักศึกษาหญิงที่กำลังศึกษาในระดับปริญญาตรีชั้นปีที่ 4 เพื่อเป็นตัวบ่งชี้สัดส่วนของแรงงานหญิงที่จบการศึกษาระดับปริญญาตรีในอุปทานแรงงานของประเทศไทย เนื่องจากไม่มีข้อมูลที่สามารถบ่งชี้ถึงสัดส่วนแรงงานหญิงที่แท้จริงในอุปทานแรงงานของประเทศไทยและข้อมูลผู้จบการศึกษาในระดับอุดมศึกษาจากสำนักงานคณะกรรมการอุดมศึกษาไม่มีความสมบูรณ์ในหลายปีการศึกษา

คล้ายคลึงกันในทุกกลุ่มตัวอย่าง ถึงแม้จะแยกพิจารณาตามระดับการศึกษาสูงสุดของกลุ่มตัวอย่างแล้ว ซึ่งความห่างของระยะเวลาในการแก้ไขข้อมูลเหล่านี้บ่งชี้ว่า โดยปรกติแล้วผู้ใช้บริการส่วนใหญ่ ไม่ได้เข้ามาปรับปรุงข้อมูลประวัติการทำงานของตนเองในฐานะข้อมูลของ IJP อย่างสม่ำเสมอ ซึ่งพฤติกรรมดังกล่าวเป็นหนึ่งในข้อจำกัดของข้อมูล ORE นี้ ที่ข้อมูลของตัวอย่างบางตัวอย่างอาจจะไม่สามารถนำมาใช้ในการวิเคราะห์ได้ เช่น หากตัวอย่างทำการฝากประวัติโดยทันทีหลังจากจบการศึกษา แต่ไม่ได้เข้ามาทำการปรับปรุงประวัติหลังจากได้งานทำแล้ว ทำให้ข้อมูลที่ปรากฏในฐานะข้อมูลมีเพียงข้อมูลที่เกี่ยวข้องกับการศึกษาของตัวอย่าง ซึ่งทำให้เกิดข้อจำกัดกับตัวอย่างนี้ หากผู้วิจัยต้องการนำไปวิเคราะห์ในประเด็นที่ต้องใช้ข้อมูลค่าจ้าง/เงินเดือนจากการทำงาน อย่างเช่น ผลตอบแทนทางการศึกษา เป็นต้น ตารางที่ 1 ยังชี้ให้เห็นถึงความสัมพันธ์ที่สามารถคาดหมายได้ระหว่าง อายุ (ในที่นี้คือ อายุ ณ วันที่มีการแก้ไขข้อมูลครั้งล่าสุด) ข้อมูลประสบการณ์การทำงาน และเงินเดือนขั้นต่ำที่ต้องการ ซึ่งโดยเฉลี่ยแล้ว กลุ่มตัวอย่างที่จบการศึกษาในระดับปริญญาเอกจะมีอายุ (ณ วันที่แก้ไขข้อมูลล่าสุด) เท่ากับ 41 ปี ซึ่งสูงที่สุดเมื่อเทียบกับผู้จบการศึกษาในระดับการศึกษาอื่นๆ ทั้งยังเป็นกลุ่มตัวอย่างที่มีประสบการณ์การทำงานและเงินเดือนที่คาดหวังสูงสุด คือ 92 เดือน และ 43,217 บาทต่อเดือน ตามลำดับ โดยกลุ่มตัวอย่างที่มีอายุ (ณ วันที่แก้ไขข้อมูลล่าสุด) น้อยที่สุด คือ กลุ่มตัวอย่างที่จบการศึกษาในระดับมัธยมศึกษาตอนต้น (ม.3) โดยมีอายุเท่ากับ 26 ปี และยังเป็นกลุ่มตัวอย่างที่มีประสบการณ์การทำงาน (40 เดือน) และเงินเดือนที่คาดหวังน้อยที่สุด (10,289 บาทต่อเดือน) เช่นเดียวกัน Appendix A แสดงรายละเอียดเพิ่มเติมสำหรับคุณลักษณะของข้อมูล ORE ในกลุ่มตัวอย่างผู้จบการศึกษาใหม่

โดยถึงแม้ว่า ตารางที่ 1 จะแสดงให้เห็นว่า “โดยเฉลี่ย” แล้ว กลุ่มตัวอย่างแรงงานที่ใช้บริการฝากประวัติการทำงานกับ IJP นั้น อาจจะไม่ได้ทำการปรับปรุงประวัติการทำงานบ่อยครั้งนัก (หากพิจารณาถึงความแตกต่างระหว่างอายุ ณ ปี พ.ศ. 2562 และ อายุ ณ วันที่มีการแก้ไขข้อมูลล่าสุด) ซึ่งอาจจะทำให้เข้าใจว่าข้อมูล ORE เป็นข้อมูลที่ไม่เป็นปัจจุบัน ตารางที่ 2 จึงทำการเปรียบเทียบจำนวนตัวอย่างที่มีประสบการณ์การทำงานของข้อมูล ORE และจำนวนตัวอย่าง “ลูกจ้างเอกชน” ของข้อมูล LFS เพื่อแสดงพลวัตของการปรับปรุงข้อมูลประวัติการทำงาน โดยใช้ข้อมูลวันเดือนปีที่ปรากฏอยู่ในข้อมูลการแก้ไขประวัติการทำงานครั้งล่าสุดมาเป็นตัวแทนของช่วงเวลาที่มีข้อมูล ORE ปรากฏขึ้นมาในฐานะข้อมูล เนื่องจากเว็บไซต์ผู้ให้บริการจัดหางานทั้งสองราย (IJP-A และ IJP-B) ทำการแสดงผลเฉพาะประวัติการทำงานออนไลน์ที่มาจากการแก้ไขครั้งล่าสุดเท่านั้น ซึ่งทำให้จำนวนตัวอย่างที่ได้เป็นจำนวนตัวอย่างที่ไม่เคยถูกนับมาก่อนและปราศจากปัญหาการนับซ้ำ

จำนวนตัวอย่างที่มีการปรับปรุงข้อมูลประวัติการทำงานในแต่ละปี ดังแสดงใน ตารางที่ 2 ชี้ให้เห็นว่ากลุ่มตัวอย่างแรงงานที่เข้ามาใช้บริการรับฝากประวัติ มักจะทำการปรับปรุงข้อมูลประวัติในช่วงไตรมาสที่ 4 ของแต่ละปีมากที่สุด รองลงมาคือ ไตรมาสที่ 2 และ ไตรมาสที่ 3 ตามลำดับ โดยที่ไตรมาสที่ 1 คือ ไตรมาสที่กลุ่มตัวอย่างแรงงานเลือกที่จะปรับปรุงประวัติน้อยที่สุด และรูปแบบของการปรับปรุงข้อมูลประวัติการทำงานดังกล่าวนี้ มีรูปแบบการเกิดซ้ำเป็นประจำทุกปี⁸ ซึ่งหากรูปแบบพฤติกรรมการปรับปรุงข้อมูลประวัติสามารถใช้เป็นตัวบ่งชี้ถึงระดับความต้องการแรงงานในขณะนั้นได้ หนึ่งในประเด็นที่น่าสนใจสำหรับการศึกษาเพิ่มเติม ก็คือ การศึกษาผลกระทบของการจบการศึกษาระดับปริญญาตรีภายในสามปีครึ่ง เนื่องจาก ช่วงเวลาที่แรงงานใหม่ที่จบการศึกษาภายในสามปีครึ่งเข้าสู่ตลาดแรงงานนั้น มักจะอยู่ในช่วงเดือนมกราคม ซึ่งหากลักษณะของพลวัตในตลาดแรงงานเป็นไปตามข้อมูลใน ตารางที่ 2 แสดงว่า แรงงานใหม่ที่เพิ่งจบการศึกษาเหล่านี้ จะเข้ามาในช่วงที่ระดับความต้องการแรงงานลดลง ซึ่งอาจส่งผลให้แรงงานใหม่ต้องรอคอยงาน และอาจก่อให้เกิดผลกระทบในระยะยาวได้ (scarring effects)

⁸ สาเหตุที่ข้อมูลจำนวนตัวอย่างที่มีการปรับปรุงข้อมูลประวัติใน ตารางที่ 2 เริ่มแสดงจาก พ.ศ. 2552 เนื่องจากข้อมูลการเข้าชมเว็บไซต์ของผู้ให้บริการทั้งสอง บ่งชี้ว่าปีดังกล่าวน่าจะเป็นปีแรกๆของการดำเนินการของเว็บไซต์

ตารางที่ 1: คุณลักษณะของข้อมูล ORE ในกลุ่มตัวอย่างที่มีประสบการณ์ทำงาน

คุณลักษณะ	รวมทุกระดับ	ป.เอก	ป.โท	ป.ตรี	ปวส.
เพศหญิง (%)	63.29	46.74	61.62	67.87	52.86
สถานภาพสมรส: โสด (%)	65.68	51.09	61.72	68.87	59.98
สถานภาพสมรส: สมรส (%)	12.25	23.91	144.33	10.27	15.91
อายุ: ณ ปี 2562 (ปี)	32.41	44.34	37.61	32.63	32.51
อายุ: ณ วันที่แก้ไขข้อมูลล่าสุด (ปี)	28.11	41.32	33.05	28.25	28.16
อายุ: ณ ปีที่จบการศึกษา (ปี)	22.58	38.36	28.94	23.60	21.99
ประสบการณ์การทำงาน (เดือน)	42.97	92.35	64.38	40.39	49.32
เงินเดือนที่คาดหวัง (บาท/เดือน)	14,443	43,217	26,398	15,700	12,115
จำนวนตัวอย่าง	528,725	92	11,227	333,775	91,729

คุณลักษณะ	รวมทุกระดับ	ปวช.	ม.6	ม.3	ป.6
เพศหญิง (%)	63.29	59.29	59.48	48.29	27.63
สถานภาพสมรส: โสด (%)	65.68	59.35	62.17	57.01	54.03
สถานภาพสมรส: สมรส (%)	12.25	16.20	14.42	16.78	22.66
อายุ: ณ ปี 2562 (ปี)	32.41	31.15	30.86	30.01	32.33
อายุ: ณ วันที่แก้ไขข้อมูลล่าสุด (ปี)	28.11	27.14	27.00	26.19	28.76
อายุ: ณ ปีที่จบการศึกษา (ปี)	22.58	19.27	19.24	16.11	13.33
ประสบการณ์การทำงาน (เดือน)	42.97	45.82	42.42	39.69	52.12
เงินเดือนที่คาดหวัง (บาท/เดือน)	14,443	11,020	10,601	10,289	11,105
จำนวนตัวอย่าง	528,725	32,041	44,792	13,679	1,390

หมายเหตุ:

- กลุ่มตัวอย่างที่มีประสบการณ์ทำงานที่นำมาใช้ในตารางนี้ เลือกใช้เฉพาะ
 - ตัวอย่างที่มีข้อมูลเงินเดือนที่คาดหวัง และ
 - ตัวอย่างที่ไม่พบสิ่งบ่งชี้ความผิดปกติในข้อมูลเงินเดือนที่คาดหวัง คือ (1) ตัวอย่างที่มีเงินเดือนที่ด้อยกว่า 3,000 บาทต่อเดือน หรือ (2) ตัวอย่างที่มีเงินเดือนที่มีประสบการณ์การทำงานน้อยกว่า 120 เดือน (5 ปี) และจบการศึกษาสูงสุดในระดับ ประถมศึกษา/มัธยมศึกษาตอนต้น/มัธยมศึกษาตอนปลาย/ปวช./ปวส. แต่ระบุเงินเดือนที่คาดหวังมากกว่า 50,000 บาทต่อเดือน และ
 - ตัวอย่างที่ไม่พบความผิดปกติในข้อมูลระยะเวลาของประสบการณ์ทำงาน (จำนวนเดือน) โดยทำการคำนวณประสบการณ์ทำงานที่น่าจะเป็นไปได้จากอายุ ณ วันที่แก้ไขข้อมูลล่าสุดเทียบกับอายุที่น้อยที่สุดที่เป็นไปได้ที่สามารถจบการศึกษาในแต่ละระดับ โดยประสบการณ์ทำงานที่ระบุอยู่ใน ORE จะต้องมากกว่าประสบการณ์ทำงานที่คำนวณจากอายุ ณ วันที่แก้ไขข้อมูลล่าสุดไม่เกิน 12 เดือน ดังนั้น จำนวนตัวอย่างทั้งหมดที่ใช้ในตารางนี้เหลือเพียง 528,725 คน จากจำนวนกลุ่มตัวอย่างที่มีประสบการณ์ทำงานทั้งหมด 710,788 คน ในเบื้องต้น
- คุณลักษณะเพศหญิงและสถานภาพสมรส แสดงข้อมูลเป็นสัดส่วน (%) ส่วนคุณลักษณะอายุ ประสบการณ์การทำงาน และ เงินเดือนที่คาดหวัง แสดงข้อมูลเป็นค่าเฉลี่ย

ตารางที่ 2: เปรียบเทียบจำนวนตัวอย่างที่มีประสิทธิภาพการทำงานของข้อมูล ORE และจำนวนตัวอย่าง “ลูกจ้างเอกชน” ของข้อมูล LFS (หน่วย: คน)

ปี พ.ศ.	ไตรมาส 1		ไตรมาส 2		ไตรมาส 3		ไตรมาส 4	
	ORE	LFS	ORE	LFS	ORE	LFS	ORE	LFS
2552	120	39894	339	39112	588	37553	1503	38920
2553	11753	40189	17289	39157	14843	25404	11547	37576
2554	9231	39214	15680	37530	13528	37022	10670	35966
2555	8970	40135	15030	39878	13640	40359	10993	39597
2556	8623	39658	15098	39005	13199	35984	56313	35381
2557	8105	34880	13509	34783	13574	34870	21338	34134
2558	8410	34648	15747	34983	14180	34273	23627	33618
2559	9022	34824	17125	34942	14762	34929	22075	32514
2560	8902	33798	17380	34527	15577	33847	21259	32016
2561	9825	30134	17866	33627	16473	33760	24210	32482

หมายเหตุ:

- จำนวนตัวอย่างของชุดข้อมูล ORE ในแต่ละไตรมาสของแต่ละปี คำนวณจากชุดข้อมูล ORE โดยใช้วันเดือนปีที่ปรากฏอยู่ในข้อมูลการแก้ไขประวัติการทำงานครั้งล่าสุด มาเป็นตัวแทนของช่วงเวลาที่ข้อมูล ORE ปรากฏขึ้นมาในฐานข้อมูล เนื่องจากเว็บไซต์ผู้ให้บริการจัดหางานทั้งสองรายทำการแสดงเฉพาะประวัติการทำงานออนไลน์ที่มาจากการแก้ไขครั้งล่าสุดเท่านั้น ซึ่งทำให้จำนวนตัวอย่างที่ได้ เป็นจำนวนตัวอย่างที่ไม่เคยถูกนับมาก่อน ปรมาจากปัญหาการนับซ้ำ นอกจากนี้ การคำนวณจะใช้เฉพาะตัวอย่างที่มีข้อมูลประวัติการทำงานและปรากฏข้อมูลเงินเดือนขั้นต่ำที่ต้องการ เท่านั้น
- จำนวนตัวอย่างของ LFS ในแต่ละไตรมาสในแต่ละปี คำนวณจาก microdata ของชุดข้อมูล LFS โดยกำหนดให้สถานะทำงานให้เป็น “ลูกจ้างเอกชน” (WKSTAT = 6) และลักษณะการทำงานเป็นผู้ที่มีงานทำที่ยังได้รับค่าจ้าง/เงินเดือน (RE_WK = {1,2})

หนึ่งข้อสังเกตหนึ่งที่น่าสนใจจาก **ตารางที่ 1** และ อาจสอดคล้องกับนโยบายทางการศึกษาที่หลายรัฐบาลที่ผ่านมามีพยายามส่งเสริมให้มีจำนวนนักศึกษาในสายอาชีพศึกษาเพิ่มขึ้น ก็คือ การที่เงินเดือนที่คาดหวังโดยเฉลี่ยของกลุ่มตัวอย่างที่จบการศึกษาสูงสุดในระดับประกาศ นียบัตรวิชาชีพ (ปวช.) นั้น สูงกว่าเงินเดือนที่คาดหวังโดยเฉลี่ยของผู้จบการศึกษาในระดับมัธยมศึกษาตอนปลาย ถึงแม้ว่าคุณลักษณะต่างๆ ของกลุ่มตัวอย่างทั้งสองกลุ่ม ไม่ว่าจะเป็นส่วนเพศหญิง สัดส่วนสถานภาพสมรส อายุ และประสิทธิภาพการทำงาน จะมีความคล้ายคลึงกันเป็นอย่างมาก ทั้งนี้ หากเราใช้ข้อสมมติฐานว่าการพิจารณาตั้งระดับเงินเดือนที่คาดหวังในข้อมูลประวัตินั้น กลุ่มตัวอย่างเป็นผู้มีเหตุผลและตั้งระดับเงินเดือนโดยพิจารณาจากความสามารถ ประสิทธิภาพ และเงินเดือนที่ได้รับล่าสุดหรือปัจจุบัน ซึ่งหากสมมติฐานดังกล่าวเป็นจริงแสดงว่า โดยเฉลี่ยแล้วกลุ่มตัวอย่างที่จบการศึกษาระดับ ปวช. ในข้อมูล ORE นี้ อาจจะได้รับเงินเดือนที่สูงกว่ากลุ่มตัวอย่างที่จบการศึกษาระดับ ม.6

อย่างไรก็ตาม ถึงแม้ว่าข้อสมมติฐานจะเป็นจริง แต่ข้อสรุปดังกล่าวก็อาจเกิดจากปัญหา selection bias จากการใช้บริษัทหรือองค์กรที่ใช้บริการ IJP เพื่อเป็นช่องทางในการประกาศรับสมัครงานนั้น เป็นบริษัทหรือองค์กรที่มีความต้องการ

ใช้ทักษะของคนที่จบการศึกษาในสาขาชีวศึกษามากกว่าคนที่จบการศึกษาในสายสามัญ ซึ่งเสมือนเป็นการดึงคุณค่าให้แรงงานที่จบการศึกษาในสาขาชีวศึกษาเฉพาะ เข้ามาใช้บริการฝากข้อมูลประวัติ ในขณะที่แรงงานที่อาจจะมีทักษะที่ดีแต่จบการศึกษาในสายสามัญ อาจเลือกหางานโดยใช้ช่องทางอื่นมากกว่า ส่วนที่ 4 จะทำการอภิปรายเพิ่มเติมถึงความเป็นไปได้ของปัญหาของการกระจุกตัวของกลุ่มตัวอย่างจากบางกลุ่มสาขาวิชาของข้อมูล ORE

3. Stylized Facts ของโครงสร้างอุปทานแรงงานไทยจากข้อมูล ORE

ด้วยจำนวนตัวอย่างของผู้ที่จบการศึกษาสูงสุดในระดับปริญญาตรีเป็นจำนวนมาก ในส่วนนี้จึงจะใช้ประโยชน์จากข้อมูลดังกล่าว เพื่อทำการอภิปรายคุณลักษณะที่โดดเด่น (stylized facts) ของกลุ่มตัวอย่างที่จบการศึกษาสูงสุดในระดับปริญญาตรีโดยเฉพาะ เพื่อฉายภาพที่น่าสนใจของโครงสร้างอุปทานแรงงานของประเทศไทย ทั้งนี้ (ยกเว้น stylized fact 4) เพื่อให้เห็นภาพพลวัตหรือการเปลี่ยนแปลงในแต่ละช่วงเวลา การอภิปรายในส่วนนี้จึงใช้การแบ่งกลุ่มตัวอย่างตามช่วงปีเกิดหรือการแบ่งเป็น “เจนเนอเรชัน” ซึ่งงานวิจัยนี้ใช้การแบ่งเจนเนอเรชันตามคำนิยามสำนักงานสถิติแห่งชาติ ดังนี้

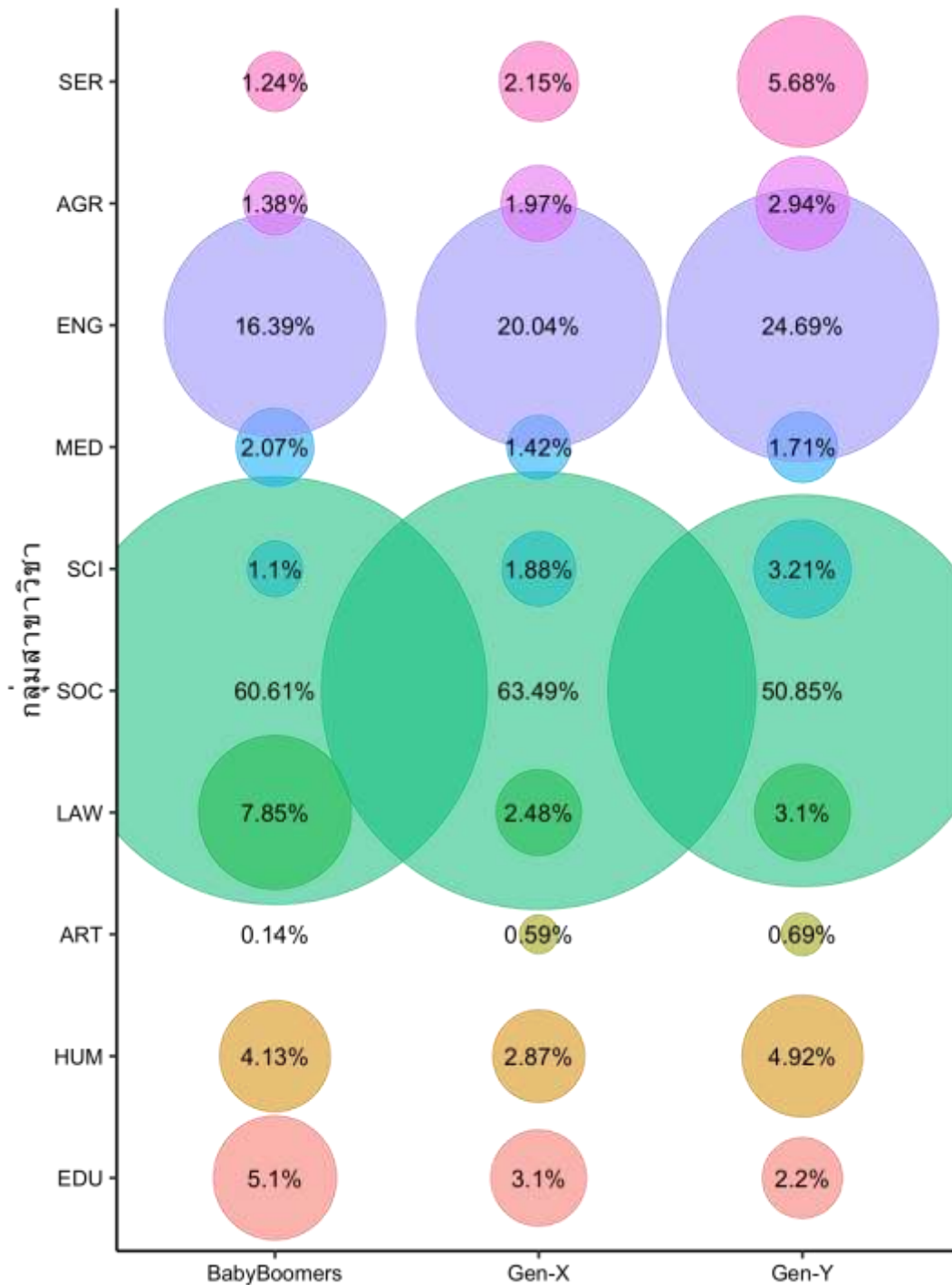
- Baby Boomers หมายถึง ผู้ที่เกิดในช่วงปี พ.ศ. 2489-2507
- Gen-X หมายถึง ผู้ที่เกิดในช่วงปี พ.ศ. 2508-2522
- Gen-Y หมายถึง ผู้ที่เกิดในช่วงปี พ.ศ. 2523-2540

Stylized fact 1: เจนเนอเรชันเปลี่ยน รูปแบบการเลือกสาขาวิชาไม่เปลี่ยน

กลุ่มสาขาวิชาที่เป็นนิคม และมักถูกเลือกในลำดับหนึ่งและสองในทั้งสามเจนเนอเรชันมีความคล้ายคลึงกัน คือ กลุ่มสาขาวิชาสังคมศาสตร์และบริหารธุรกิจ และ กลุ่มสาขาวิชาวิศวกรรมศาสตร์ ตามลำดับ รูปที่ 3 แสดงสัดส่วนของตัวอย่างที่จบการศึกษาสูงสุดในระดับปริญญาตรีของข้อมูล ORE แยกตามกลุ่มสาขาวิชาและเจนเนอเรชัน อย่างไรก็ตาม สัดส่วนการเลือกเรียนในสาขาวิชาสังคมศาสตร์และบริหารธุรกิจของ Gen-Y ก็ลดลงเป็นอย่างมากเมื่อเทียบกับกลุ่ม Baby Boomers และ Gen-X ในขณะที่สัดส่วนการเลือกเรียนในสาขาวิชาวิศวกรรมศาสตร์ก็มีการเพิ่มขึ้นอย่างต่อเนื่อง

สิ่งที่น่าสนใจ โดยเฉพาะอย่างยิ่งสำหรับ Gen-Y ซึ่งเป็นกลุ่มที่มีการใช้เทคโนโลยีสูงที่สุดในทั้งสามกลุ่ม ก็คือ แนวโน้มการเลือกเรียนในสาขาวิชาที่เกี่ยวกับเทคโนโลยี ที่นอกเหนือไปจากสาขาวิศวกรรมศาสตร์ อย่างเช่น สาขาวิชาวิทยาศาสตร์กายภาพ ซึ่งอาจมีสาเหตุมาจากการที่มีสาขาวิชาและ Keywords ที่เกี่ยวข้องกับเทคโนโลยีที่งานวิจัยนี้ได้ใช้ในการจัดกลุ่มสาขาวิชารวมอยู่เป็นจำนวนมาก (ดูรายละเอียดของ Keywords เพิ่มเติมใน Appendix A) สาขาวิชาที่เกี่ยวข้องกับงานบริการอย่างสาขาวิชาบริการ โรงแรมและการท่องเที่ยว ก็เป็นอีกหนึ่งสาขาที่คนในกลุ่ม Gen-Y สนใจและเลือกที่จะเรียนเพิ่มเป็นจำนวนมากเมื่อเทียบกับคนในกลุ่ม Baby Boomers และ Gen-X นอกจากนั้น ถึงแม้ว่าสัดส่วนที่เพิ่มขึ้นมาจะไม่มากนัก แต่การเพิ่มขึ้นของสัดส่วนของคนที่ยื่นในสาขาวิชาเกษตรศาสตร์ของคนในกลุ่ม Gen-Y ซึ่งเป็นกลุ่มคนที่ (โดยนิยาม) น่าจะชอบความสะดวกสบาย ก็เป็นเรื่องที่เกินความคาดหมายสำหรับคนเจนเนอเรชันนี้ ซึ่งเหตุผลหลักที่น่าจะมาจากความต้องการในตลาดแรงงานในธุรกิจที่เกี่ยวข้องกับการเกษตรที่เพิ่มสูงขึ้น ซึ่งเห็นได้จากสัดส่วนที่เพิ่มขึ้นเช่นเดียวกันในกลุ่มสายงานการเกษตร ประมง และ ป่าไม้ ใน Stylized fact 2

โดยสาขาวิชาอย่างศึกษาศาสตร์ ที่โดยอาชีพแล้วอาจจะต้องใช้ความอดทนสูงและต้องใช้เวลาในการเจริญก้าวหน้ารวมทั้งค่าตอบแทนโดยเฉลี่ยที่ต่ำเมื่อเทียบกับกลุ่มอาชีพอื่น ก็อาจเป็นสาเหตุให้ความนิยมเรียนในสาขานี้ลดลง และถูกเลือกเรียนน้อยลงอย่างชัดเจนและต่อเนื่องจากคนในกลุ่ม Gen-X จนมาถึง Gen-Y



รูปที่ 3: สัดส่วนของตัวอย่างที่จบการศึกษาสูงสุดในระดับปริญญาตรีของข้อมูล ORE แยกตามกลุ่มสาขาวิชา

หมายเหตุ:

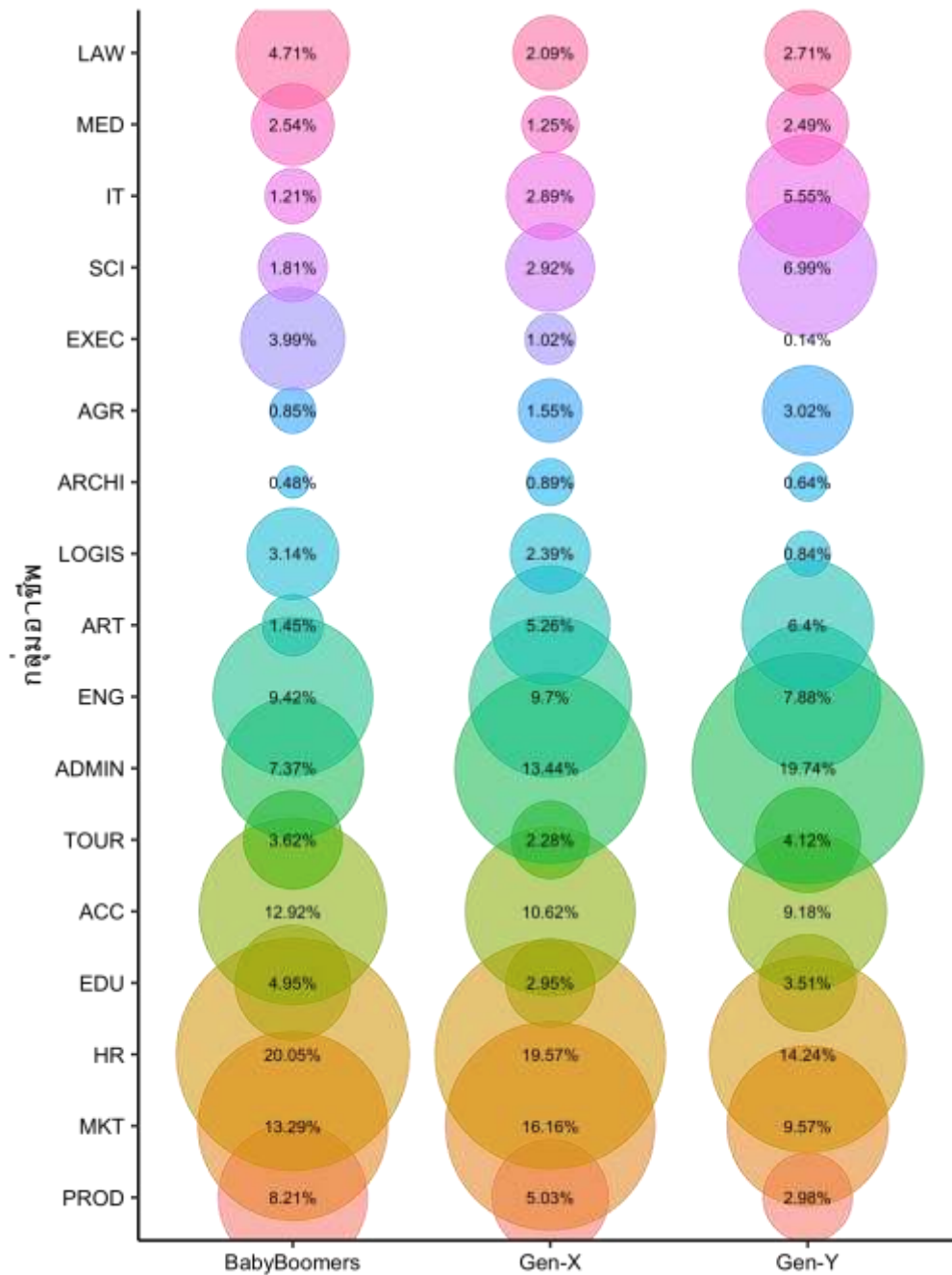
1. EDU หมายถึง สาขาวิชาศึกษาศาสตร์ HUM หมายถึง สาขาวิชามนุษยศาสตร์ ART หมายถึง สาขาวิชาศิลปกรรมศาสตร์ LAW หมายถึง สาขาวิชานิติศาสตร์ SOC หมายถึง สาขาวิชาสังคมศาสตร์และบริหารธุรกิจ SCI หมายถึง สาขาวิชาวิทยาศาสตร์กายภาพ MED หมายถึง สาขาวิชาการแพทย์และสาธารณสุข ENG หมายถึง สาขาวิชาวิศวกรรมศาสตร์ AGR หมายถึง สาขาวิชาเกษตรศาสตร์ และ SER หมายถึง สาขาวิชาบริการโรงแรมและการท่องเที่ยว
2. ดูรายละเอียดเพิ่มเติม Keywords สำหรับการจัดตัวอย่างข้อมูล ORE เข้าสู่กลุ่มสาขาวิชาใน Appendix A
3. จำนวนตัวอย่างกลุ่ม “Baby Boomers” เท่ากับ 726 ตัวอย่าง กลุ่ม “Gen-X” เท่ากับ 30,609 ตัวอย่าง และกลุ่ม “Gen-Y” เท่ากับ 253,321 ตัวอย่าง

Stylized fact 2: กลุ่มสายงานยอดนิยมยังเหมือนเดิม แตกต่างกันในสายงานที่นิยมรองลงมา

รูปที่ 4 แสดงสัดส่วนของตัวอย่างที่จบการศึกษาสูงสุดในระดับปริญญาตรีของข้อมูล ORE แยกตามกลุ่มสาขาอาชีพ เช่นเดียวกับกลุ่มสาขาวิชาอุดมศึกษา กลุ่มสายงานที่เป็นที่นิยมและคนจากทั้งสามเจนเนอเรชันเลือกที่จะประกอบอาชีพก็คือ กลุ่มอาชีพการขายและการตลาด กลุ่มอาชีพการบุคคล กลุ่มอาชีพบัญชี การเงิน ธนาคารและเศรษฐศาสตร์ และ กลุ่มอาชีพธุรการและการจัดการทั่วไป ซึ่งเป็นกลุ่มอาชีพเป้าหมายโดยตรง ของแรงงานที่จบการศึกษาในสาขาวิชาสังคมศาสตร์และบริหารธุรกิจอยู่แล้ว เช่นเดียวกับกับสาขาวิชาวิศวกรรมศาสตร์และกลุ่มอาชีพวิศวกร

การเพิ่มขึ้นอย่างชัดเจนของสัดส่วนผู้ที่ประกอบอาชีพในสายงานทางด้านคอมพิวเตอร์ ไอที และ ซอฟต์แวร์ รวมถึง กลุ่มอาชีพด้านวิทยาศาสตร์กายภาพ ของคนในกลุ่ม Gen-Y เหล่านี้เป็นที่คาดหมายได้ หากพิจารณาถึงจำนวนผู้ประกอบการ หรือ ตำแหน่งงานที่มีความจำเป็นจะต้องใช้ทักษะที่เกี่ยวข้อง เพื่อตอบสนองแนวทางการใช้ชีวิตประจำวันของคนในยุคนี้ แต่อีกสิ่งหนึ่งที่น่าสนใจและอาจเกี่ยวข้อง โดยต่อับความต้องการแรงงานที่มีทักษะทางด้านคอมพิวเตอร์และไอทีเพิ่มขึ้น ก็คือสายอาชีพที่เกี่ยวข้องกับศิลปะและสื่อสารมวลชน ซึ่งมีสัดส่วนเพิ่มสูงขึ้นเป็นอย่างมากโดยเฉพาะหากเทียบกับคนใน Baby Boomers ซึ่งมักจะถูกมองว่าเป็นคนอนุรักษ์นิยม

อีกสิ่งหนึ่งที่น่าสนใจจากรูปที่ 4 ก็คือ การลดลงอย่างต่อเนื่องของสัดส่วนผู้ประกอบอาชีพทางการผลิตและโรงงาน ซึ่งถือได้ว่าเป็นกลุ่มอาชีพที่เกี่ยวข้องกับแรงงานที่จบจากกลุ่มสาขาวิชาวิศวกรรมศาสตร์โดยตรง ซึ่งถึงแม้ว่ากลุ่มวิชาวิศวกรรมศาสตร์และงานทางด้านวิศวกรรมแบบอื่นจะยังเป็นที่สนใจสำหรับคนในกลุ่ม Gen-Y แต่เนื่องจากงานการผลิตและงานในโรงงานนั้น เป็นงานที่ต้องอยู่ในกรอบและมีลักษณะการทำงานที่เป็นขั้นตอน ตลอดจน รูปแบบเวลาทำงานที่ตายตัวและต้องทำเป็นประจำสม่ำเสมอ ก็อาจจะเป็นสาเหตุให้คนในกลุ่ม Gen-Y นี้ไม่นิยมที่จะทำหลังจากจบการศึกษา



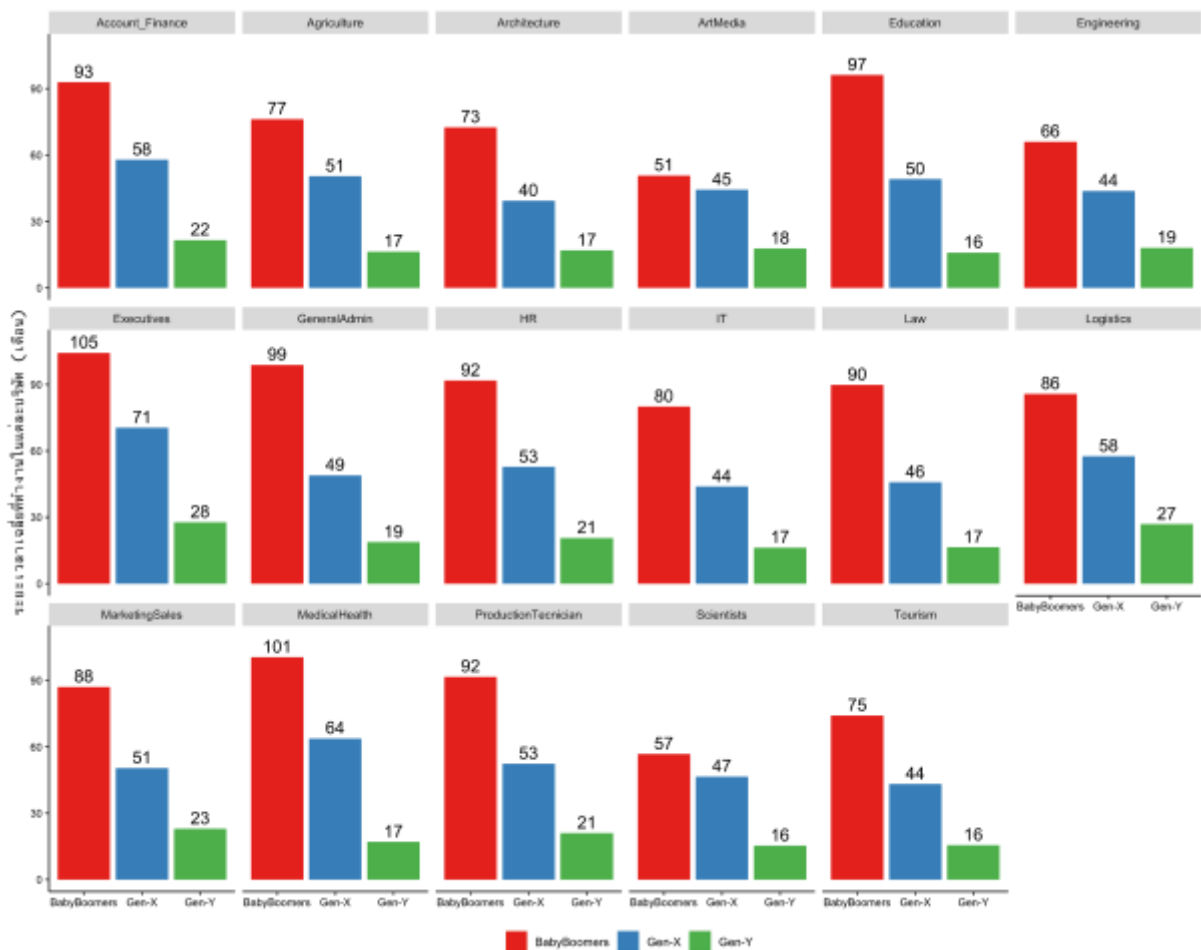
รูปที่ 4: สัดส่วนของตัวอย่างที่จบการศึกษาสูงสุดในระดับปริญญาตรีของข้อมูล ORE แยกตามกลุ่มสาขาอาชีพ

หมายเหตุ:

1. PROD หมายถึง กลุ่มอาชีพโรงงานและการผลิต MKT หมายถึง กลุ่มอาชีพการขายและการตลาด HR หมายถึง กลุ่มอาชีพการบุคคล EDU หมายถึง กลุ่มอาชีพครูศาสตร์ ACC หมายถึง กลุ่มอาชีพบัญชี การเงิน ธนาคารและเศรษฐศาสตร์ TOUR หมายถึง กลุ่มอาชีพด้าน โรงแรมและการท่องเที่ยว ADMIN หมายถึง กลุ่มอาชีพธุรการและการจัดการทั่วไป ENG หมายถึง กลุ่มอาชีพวิศวกร ART หมายถึง กลุ่มอาชีพด้านศิลปกรรมและสื่อสารมวลชน LOGIS หมายถึง กลุ่มอาชีพด้านการขนส่ง ARCHI หมายถึง กลุ่มอาชีพสถาปนิก AGR หมายถึง กลุ่มอาชีพด้านการเกษตร ประมง และป่าไม้ EXEC หมายถึง กลุ่มอาชีพในระดับผู้บริหาร SCI หมายถึง กลุ่มอาชีพด้านวิทยาศาสตร์กายภาพ IT หมายถึง กลุ่มอาชีพด้านไอที ซอฟต์แวร์และคอมพิวเตอร์ MED หมายถึง กลุ่มอาชีพแพทย์และสาธารณสุข LAW หมายถึง กลุ่มอาชีพทนายและที่ปรึกษากฎหมาย
2. ดูรายละเอียดเพิ่มเติม Keywords สำหรับการจัดตัวอย่างข้อมูล ORE เข้าสู่กลุ่มสาขาอาชีพใน Appendix B
3. จำนวนตัวอย่างกลุ่ม “Baby Boomers” เท่ากับ 828 ตัวอย่าง กลุ่ม “Gen-X” เท่ากับ 33,992 ตัวอย่าง และกลุ่ม “Gen-Y” เท่ากับ 282,474 ตัวอย่าง

Stylized fact 3: แนวโน้มการทำงานองค์กรเดียวเป็นเวลานานลดลงอย่างต่อเนื่อง

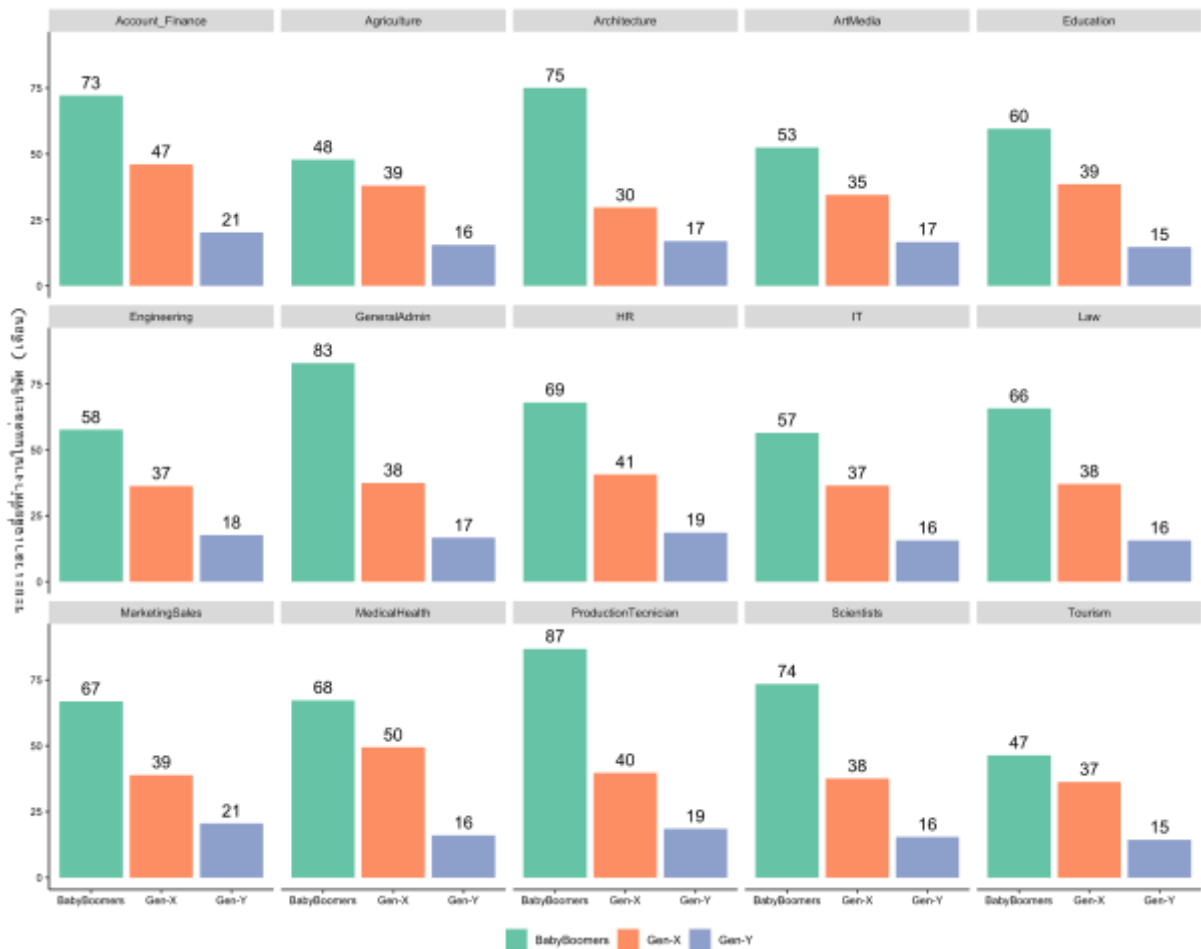
เนื่องจากจำนวนรายละเอียดของประสบการณ์ทำงาน ที่ผู้ฝากประวัติสามารถระบุได้ในข้อมูลประวัติออนไลน์ ถูกจำกัดให้ไม่เกินสามบริษัท ด้วยเหตุนี้ ผู้วิจัยจึงทำการคำนวณหาระยะเวลาที่กลุ่มตัวอย่างของข้อมูล ORE ทำงานในบริษัทแรกสำหรับกลุ่มตัวอย่างที่มีการระบุประสบการณ์ทำงานมาแล้วอย่างน้อยสองบริษัท และ ระยะเวลาที่กลุ่มตัวอย่างของข้อมูล ORE ทำงานในสองบริษัทแรก สำหรับกลุ่มตัวอย่างที่มีการระบุประสบการณ์ทำงานมาแล้วอย่างน้อยสามบริษัท เพื่อที่จะลดความเบี่ยงเบนอันเกิดจากผลของระยะเวลาในบริษัทปัจจุบันที่ตัวอย่างกำลังทำงานอยู่ ยกตัวอย่างเช่น ผู้ฝากประวัติ A เคยผ่านการทำงานที่บริษัท X และ Y มาแล้วที่ละ 6 ปี และปัจจุบันกำลังทำงานอยู่ในบริษัท Z ซึ่งถือเป็นบริษัทที่ 3 ของ A ดังนั้น ในข้อมูลประวัติออนไลน์ ผู้ฝากประวัติ A จะระบุวันสุดท้ายของการทำงานที่บริษัท Z เป็น “ปัจจุบัน” ซึ่งหากบังเอิญว่า A เพิ่งจะทำงานที่บริษัท Z ได้เพียง 1 เดือน ค่าเฉลี่ยของระยะเวลาที่ A ทำงานในแต่ละบริษัทจะเท่ากับ $(72+72+1)/3 = 50$ เดือน ซึ่งหาก A ไม่ได้มีการเปลี่ยนงานจริงและทำงานต่อไปจนกระทั่งครบ 5 ปี แล้วจึงเปลี่ยนบริษัท ค่าเฉลี่ยของระยะเวลาที่ A ทำงานในแต่ละบริษัทก็ควรจะเท่ากับ $(72+72+60)/3 = 68$ เดือน



รูปที่ 5: ระยะเวลาเฉลี่ยที่ทำงานใน “บริษัทแรกที่ระบุในข้อมูลประวัติ” ของข้อมูล ORE แยกตามกลุ่มสาขาอาชีพ

หมายเหตุ:

1. กลุ่มตัวอย่างเฉพาะผู้ที่จบการศึกษาสูงสุดในระดับปริญญาตรี
2. จำนวนตัวอย่างกลุ่ม “Baby Boomers” เท่ากับ 524 ตัวอย่าง กลุ่ม “Gen-X” เท่ากับ 22,926 ตัวอย่าง และกลุ่ม “Gen-Y” เท่ากับ 130,062 ตัวอย่าง



รูปที่ 6: ระยะเวลาในการทำงานเฉลี่ยใน “สองบริษัทแรกที่ระบุในข้อมูลประวัติ” ของข้อมูล ORE แยกตามกลุ่มสาขาอาชีพ

หมายเหตุ:

1. กลุ่มตัวอย่างเฉพาะผู้ที่จบการศึกษาสูงสุดในระดับปริญญาตรี
2. จำนวนตัวอย่างกลุ่ม “Baby Boomers” เท่ากับ 216 ตัวอย่าง กลุ่ม “Gen-X” เท่ากับ 9,493 ตัวอย่าง และกลุ่ม “Gen-Y” เท่ากับ 37,818 ตัวอย่าง
3. ไม่พบกลุ่มตัวอย่างที่อยู่ในกลุ่มอาชีพ Executives และ Logistics

รูปที่ 5 และ 6 แสดงระยะเวลาเฉลี่ยในบริษัทแรกและสองบริษัทแรกที่ระบุในข้อมูลประวัติ ของกลุ่มตัวอย่างที่จบการศึกษาสูงสุดในระดับปริญญาตรี ตามลำดับ โดยรูปทั้งสองให้ข้อสรุปที่ไม่ต่างกันในทุกกลุ่มอาชีพ (ยกเว้นกลุ่มอาชีพ Executives และ Logistics ที่ไม่พบตัวอย่างที่มีการระบุประสบการณ์ทำงานมาแล้วอย่างน้อยสามบริษัท) คือ ระยะเวลาที่ตัวอย่างทำงานโดยเฉลี่ยต่อองค์กรหรือบริษัทของ Gen-B (Baby Boomers) นานกว่า Gen-X และ ระยะเวลาที่ตัวอย่างทำงานโดยเฉลี่ยต่อองค์กรหรือบริษัทของ Gen-X ก็นานกว่า Gen-Y ซึ่งข้อสรุปที่ได้นี้ สอดคล้องเป็นอย่างมากกับคำนิยามของแต่ละเจนเนอเรชัน ที่เรามักจะเห็นได้ทั่วไปว่า⁹ นอกจากนี้ หากพิจารณาถึงระยะเวลาโดยเฉลี่ยที่ลดลงในบริษัทที่สองในรูปที่ 6 เมื่อเปรียบเทียบในเจนเนอเรชันเดียวกันและกลุ่มอาชีพเดียวกันกับรูปที่ 5 (โดยเฉพาะ Gen-B และ Gen-X) แล้ว เราจะเห็นถึง

⁹ โดยนิยามทั่วไปคน Gen-B นั้น ถูกมองว่าเป็นคนที่มีความอดทนและทุ่มเทให้กับองค์กรสูง ในขณะที่ Gen-Y นั้น ถูกมองว่าเป็นคนที่ไม่ชอบการรอคอย และต้องการประสบการณ์สำเร็จอย่างรวดเร็ว และคนที่อยู่ใน Gen-X คือคนที่มีความผูกพันอยู่ระหว่าง Gen-B และ Gen-Y

แนวโน้มที่ลดลงของระยะเวลาในการทำงานในองค์กรใดองค์กรหนึ่ง ซึ่งเป็นเครื่องบ่งชี้ถึงการปรับตัวของคนในแต่ละเจเนอเรชัน ต่อสถานการณ์ที่เปลี่ยนแปลงไปในตลาดแรงงานเมื่อเวลาผ่านไปเช่นเดียวกัน

Stylized fact 4: เงินเดือนในภาคเอกชนปรับตัวตามนโยบาย?

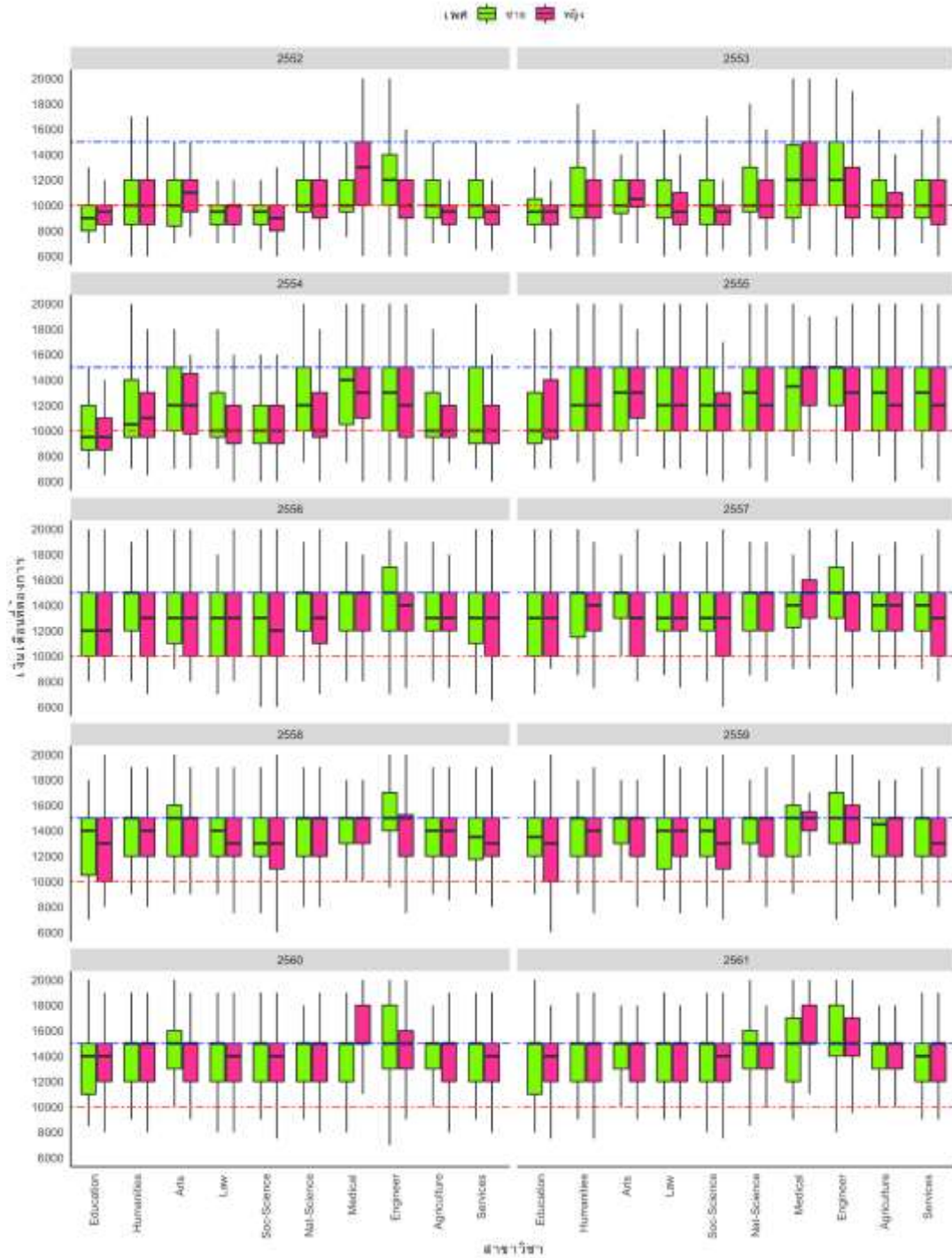
นโยบายของรัฐบาลว่าด้วยเรื่องการปรับโครงสร้างเงินเดือนข้าราชการทั้งระบบใน พ.ศ. 2555 เพื่อเพิ่มค่าครองชีพและกำลังซื้อให้กับประชาชน โดยเป็นการปรับให้ข้าราชการที่จบการศึกษาในระดับปริญญาตรีได้รับเงินเดือนขั้นต่ำ 15,000 บาท และจะมีผลบังคับใช้เฉพาะกับข้าราชการเท่านั้น (โดยมีความคาดหวังว่าการขึ้นเงินเดือนข้าราชการนี้จะส่งผลต่อเงินเดือนในภาคเอกชนด้วยเช่นกัน) ซึ่งในเบื้องต้น นโยบายดังกล่าวก็ส่งผลในทางอ้อมต่อเงินเดือนที่คาดหวังในภาคเอกชนเช่นเดียวกัน **รูปที่ 7** แสดงเงินเดือนที่คาดหวังของผู้จบการศึกษาใหม่ในระดับปริญญาตรี ของชุดข้อมูล ORE แยกตามกลุ่มสาขาวิชา ซึ่งจะเห็นได้ว่าในช่วงก่อนปี พ.ศ. 2555 นั้น เงินเดือนที่คาดหวังในแต่ละสาขาวิชามีความแตกต่างกันไป โดยกว่า 75% (Q3) ของผู้จบการศึกษาใหม่รวมทุกสาขาวิชาที่มีเงินเดือนที่คาดหวังน้อยกว่า 15,000 บาท ดังแสดงใน **ตารางที่ 3**

โดยตั้งแต่ปี พ.ศ. 2555 ซึ่งเป็นปีที่รัฐบาลให้ความเห็นชอบกฎ ก.พ. ว่าด้วยเรื่องการปรับเงินเดือนข้าราชการ เป็นต้นมา เงินเดือนที่คาดหวังของผู้จบการศึกษาใหม่ในข้อมูล ORE ซึ่งน่าจะมีความหมายในการทำงานในภาคเอกชน กลับมีการขยับตัวเข้าสู่ช่วงฐานเงินเดือน 15,000 บาท มากขึ้นเรื่อยๆ โดยจาก **ตารางที่ 3** จะเห็นได้ว่าค่ามัธยฐาน (Q2) ของเงินเดือนที่คาดหวังมีการขยับเข้าใกล้ค่าควอไทล์ที่สาม ซึ่งก็ยังมีค่าที่ 15,000 บาท มากขึ้นเรื่อยๆในแต่ละปีจนกระทั่งปี พ.ศ. 2559 และก็ไม่มีการขยับใดๆ เลยตั้งแต่นั้นเป็นต้นมา

อย่างไรก็ตาม ในความเป็นจริงนโยบายการปรับเงินเดือนข้าราชการดังกล่าว ก็อาจจะไม่มีผลกระทบที่มากนักต่อค่าจ้างเงินเดือนในภาคเอกชน ซึ่งจะเห็นได้จาก **ตารางที่ 4** ซึ่งแสดงเงินเดือนที่คาดหวังของแรงงานที่เพิ่งทำงานไม่นาน เพื่อลด upward bias ของประสบการณ์ทำงานที่อาจมีต่อค่าจ้างเงินเดือน หากใช้กลุ่มตัวอย่างที่มีประสบการณ์ทำงานสูงกว่านี้ในการวิเคราะห์ โดยใช้ข้อมูลจากกลุ่มตัวอย่างที่มีประสบการณ์ทำงานมาแล้วไม่เกิน 2 ปี รวมทุกสาขาวิชา ของข้อมูล ORE ซึ่งถึงแม้ชุดข้อมูลในงานวิจัยนี้จะมีปัญหาในเรื่องข้อมูลเงินเดือนที่แท้จริง แต่การใช้ข้อมูลเงินเดือนที่คาดหวังมาวิเคราะห์ในกรณีนี้ น่าจะเป็นตัวบ่งชี้ที่ดีถึงรายได้เงินเดือนของกลุ่มตัวอย่าง เนื่องจากมีความเป็นไปได้สูงที่เงินเดือนที่คาดหวังของกลุ่มตัวอย่างนั้น น่าจะเท่ากับหรือมากกว่าเงินเดือนที่ได้รับล่าสุดหรือปัจจุบัน

ตารางที่ 3: ควอไทล์เงินเดือนที่คาดหวังของผู้จบการศึกษาใหม่รวมทุกสาขาวิชา (หน่วย: บาทต่อเดือน)

ควอไทล์	ปีที่จบการศึกษา									
	2552	2553	2554	2555	2556	2557	2558	2559	2560	2561
Q1	8,500	9,000	9,000	10,000	10,000	12,000	12,000	12,000	12,000	12,000
Q2	10,000	10,000	10,000	12,000	13,000	13,000	14,000	15,000	15,000	15,000
Q3	12,000	12,000	13,000	15,000	15,000	15,000	15,000	15,000	15,000	15,000



รูปที่ 7: เงินเดือนที่คาดหวังของผู้ับการศึกษาใหม่ในระดับปริญญาตรี ของข้อมูล ORE แยกตามกลุ่มสาขาวิชา

หมายเหตุ:

1. นิยามของ “ผู้ับการศึกษาใหม่” ของกลุ่มตัวอย่างที่แสดงในตารางนี้ ถูกกำหนดให้ต้องมีคุณสมบัติ ดังต่อไปนี้

- ผลต่างระหว่างปีที่มีการแก้ไขข้อมูลล่าสุดและปีที่จบการศึกษาสูงสุด ต้องน้อยกว่าสองปี และ
- ไม่มีข้อมูลรายละเอียดประวัติการทำงาน (เนื่องจาก พบว่าบางตัวอย่าง ไม่ได้ระบุจำนวนปีประสบการณ์ทำงาน แต่มีการระบุรายละเอียดประวัติการทำงาน) และ
- จำนวนปีประสบการณ์ทำงานทั้งหมดต้องเท่ากับศูนย์ (เนื่องจาก พบว่าบางตัวอย่าง ไม่ได้ระบุรายละเอียดประวัติการทำงาน แต่มีการระบุจำนวนปีประสบการณ์ทำงาน)

ตารางที่ 4: ควอไทล์เงินเดือนที่คาดหวังของผู้มีประสบการณ์ทำงานไม่เกิน 2 ปี รวมทุกสาขาวิชา (หน่วย: บาทต่อเดือน)

ควอไทล์	ปี พ.ศ.									
	2552	2553	2554	2555	2556	2557	2558	2559	2560	2561
Q1	9,000	9,000	10,000	10,000	11,000	12,000	12,000	12,000	12,000	13,000
Q2	10,000	10,000	12,000	13,000	13,000	14,000	15,000	15,000	15,000	15,000
Q3	12,000	13,000	15,000	15,000	15,000	15,000	15,000	15,000	16,000	16,000

ข้อมูลใน **ตารางที่ 4** บ่งชี้ว่า มีความเป็นไปได้ที่ส่วนใหญ่ของกลุ่มตัวอย่างของข้อมูล ORE ที่เพิ่งทำงานมาไม่นาน เหล่านี้ได้รับเงินเดือนไม่ถึง 15,000 บาท อันจะเห็นได้จากการที่ค่าครึ่งหนึ่งของกลุ่มตัวอย่างมีเงินเดือนที่คาดหวังเพียง 13,000-14,000 บาท ในช่วงปี พ.ศ. 2555-2557 และตั้งแต่ปี พ.ศ. 2558 เป็นต้นมา เงินเดือนที่คาดหวังของกลุ่มตัวอย่างกว่า 50% ก็หยุดนิ่งอยู่ที่ 15,000 บาท ซึ่งสะท้อนให้เห็นถึงแนวทางการวางแผนนโยบายในอนาคต ที่อาจจะต้องใช้นโยบายที่มีการบังคับใช้กับภาคเอกชนโดยตรง แทนการบังคับใช้อ้อมเหมือนเช่นในกรณีนี้ หากรัฐบาลต้องการเพิ่มค่าครองชีพและกำลังซื้อให้กับประชาชน (ดังเช่นที่นโยบายการปรับเงินเดือนข้าราชการได้ให้เหตุผลไว้)

4. ข้อควรระวังในการใช้งานข้อมูล ORE

ปัญหาการเป็นตัวแทนประชากรถือเป็นปัญหาพื้นฐานของงานวิจัยที่ใช้ข้อมูลทุกประเภท ไม่ว่าจะข้อมูลนั้นจะเป็นข้อมูลที่เกิดจากการสังเกต (observational data) หรือข้อมูลจากการสำรวจ (survey data) ซึ่งนักวิจัยจำเป็นต้องใช้ความระมัดระวังในการวิเคราะห์ โดยเฉพาะอย่างยิ่งกับการนำไปใช้กับการศึกษาที่เป็นเชิงสาเหตุ (causal inference studies) ซึ่งปัญหาดังกล่าวก็มีแนวโน้มที่จะเกิดขึ้นกับข้อมูล ORE ที่งานวิจัยนี้ได้จัดทำขึ้นก็เช่นกัน การอภิปรายในหลายส่วนที่ผ่านมาชี้ให้เห็นถึงแนวโน้มของการที่ข้อมูล ORE จะไม่สามารถใช้เป็นตัวแทนประชากรได้ งานวิจัยที่ผ่านมามักตรวจสอบการเป็นตัวแทนประชากรของชุดข้อมูลที่ได้จากการเก็บข้อมูลทางอินเทอร์เน็ต ด้วยการเปรียบเทียบกับข้อมูลที่ได้มาจากการสำรวจขนาดใหญ่และครอบคลุมทั้งประเทศ อย่างเช่น ชุดข้อมูลการสำรวจภาวะการทำงานของประชากร

เนื่องจากการตรวจสอบปัญหาการเป็นตัวแทนประชากรด้วยการเปรียบเทียบกับข้อมูล LFS อาจไม่ใช่สิ่งที่ถูกต้องนัก ดังนั้น การอภิปรายในส่วนนี้นอกจากจะมีจุดประสงค์ที่จะชี้ให้เห็นถึงปัญหาการเป็นตัวแทนประชากรของข้อมูล ORE แล้ว อีกหนึ่งประเด็นที่ตรงการจะชี้ให้เห็น ก็คือ ประเด็นปัญหาที่เกี่ยวข้องกับวิธีการออกแบบการสำรวจ ซึ่งทำให้เกิดความเบี่ยงเบน (bias) ในค่าถ่วงน้ำหนัก ซึ่งใช้ในการคำนวณหา “ค่าประมาณ” ในระดับรวม (aggregate level) จากข้อมูล LFS ต่อไป เนื่องจากนักวิเคราะห์หรือนักวิจัยที่ใช้ชุดข้อมูล LFS มักจะมีสมมติฐานหรือความเชื่อที่ว่าชุดข้อมูล LFS สามารถเป็นตัวแทนประชากรของแรงงานไทยได้¹⁰ โดยงานวิจัยนี้ทำการตรวจสอบความตรงภายนอก (external validity) ของ “ค่าประมาณ” ในระดับรวม ที่คำนวณโดยใช้ microdata ของชุดข้อมูล LFS กับข้อมูลสถิติที่ได้จากกลุ่มตัวอย่างที่มีขนาดครอบคลุมประชากรทั้งหมดหรือเกือบทั้งหมด (universal/near-universal population coverage) อย่างเช่น (1) ข้อมูลสถิติงาน

¹⁰ Footnote 3 ของ วีระชาติ และ กิตติพงษ์ (2017) คือหนึ่งในตัวอย่างของการใช้สมมติฐานดังกล่าว โดยกล่าวไว้อย่างชัดเจนว่า “LFS เป็นข้อมูลแบบสุ่มทั่วประเทศและมีขนาดใหญ่ ทำให้ลดปัญหาอันเกิดจากความคลาดเคลื่อนของข้อมูลได้”

ประกันสังคม ซึ่งเป็นผลผลิตจากชุดข้อมูลผู้ประกันตนรายบุคคลของสำนักงานประกันสังคม ซึ่งถือได้ว่าเป็นชุดข้อมูลที่มีความครอบคลุมมากที่สุดในประเภทแรงงานภาคเอกชนที่เป็นการจ้างงานอย่างเป็นทางการ (formal employment) ของประเทศไทย และ (2) สถิติจำนวนผู้จบการศึกษาในระดับชั้นต่างๆ จากหน่วยงานที่เป็นผู้จัดเก็บข้อมูลและเผยแพร่สถิติเหล่านี้ ทั้งโดยตรงและโดยอ้อม อย่างเช่น กระทรวงศึกษาธิการ กระทรวงแรงงาน และสำนักงานคณะกรรมการการอุดมศึกษา เป็นต้น

ข้อมูลการสำรวจภาวะการทำงานของประชากร ดำเนินการสำรวจโดยสำนักงานสถิติแห่งชาติ ชุดข้อมูล LFS ถือได้ว่าเป็นชุดข้อมูลภาคตัดขวาง (cross-sectional data) จากการสำรวจ (survey-based) ที่มีกลุ่มตัวอย่างขนาดใหญ่และมีความถี่ในการสำรวจมากที่สุดของประเทศไทย¹¹ ซึ่งทำให้ข้อมูล LFS จัดอยู่ในข้อมูลประเภทตัดขวางแบบเก็บซ้ำ (repeated cross-sectional data) โดยผลผลิตที่ได้จากการคำนวณที่ใช้ข้อมูล LFS เช่น จำนวนผู้มีงานทำและจำนวนผู้ว่างงาน มักจะถูกนำไปใช้อ้างอิง การกำหนดนโยบายทางเศรษฐกิจ รวมถึงการนำไปใช้ในงานวิจัยอย่างแพร่หลาย

ชุดข้อมูล LFS ถือได้ว่าเป็นชุดข้อมูลที่บรรจุข้อมูลคุณลักษณะที่เกี่ยวข้องกับการศึกษาและการทำงานที่ละเอียด เช่น ระดับการศึกษาสูงสุด สาขาวิชา จำนวนปีของหลักสูตร และสาขาวิชา (สายสามัญ สายอาชีวศึกษา และสาขาวิชาการศึกษา) อาชีพ กลุ่มอุตสาหกรรม ฐานะในการทำงาน ขนาดกิจการ จำนวนชั่วโมงในการทำงาน เงินเดือน/ค่าจ้าง โบนัส ค่าล่วงเวลา เป็นต้น นอกจากนี้ ด้วยการใช้ชุดข้อมูล LFS ดำเนินการสำรวจสมาชิกของทั้งครัวเรือน ทำให้นักวิจัยสามารถใช้ข้อมูลของสมาชิกคนอื่นๆ ในครัวเรือน เพื่อนำมาใช้โดยตรงหรือนำมาสร้างเป็นตัวแปรใหม่เพื่อใช้บ่งบอกถึงฐานะทางเศรษฐกิจและสังคม (socio-economic status) ของครอบครัวของกลุ่มตัวอย่าง ซึ่งเป็นข้อมูลที่มีความสำคัญเป็นอย่างมากในการวิจัยที่เกี่ยวข้องกับการศึกษา (Acemoglu and Pischke, 2001; Meghir and Palme, 2005) และแรงงาน (Juhn and Murphy, 1997) อย่างไรก็ตาม ข้อมูลจากการสำรวจจำเป็นต้องใช้เวลาในการเตรียมการ การสำรวจ ตลอดจนการรายงานผล ซึ่งส่งผลให้เกิดปัญหาเรื่องความล่าช้าและอาจไม่ทันต่อสถานการณ์ที่เปลี่ยนแปลงไปอย่างรวดเร็วในโลกปัจจุบัน นอกจากนี้ ข้อมูลจากการสำรวจก็มักจะพบปัญหาข้อมูลขาดหายหรือข้อมูลคลาดเคลื่อน โดยเฉพาะกับประเด็นที่อาจจะอ่อนไหวต่อความรู้สึกของผู้ตอบแบบสำรวจ เช่น ข้อมูลการว่างงาน เป็นต้น¹²

เนื่องจากจุดประสงค์สูงสุดของการเก็บข้อมูลโดยการสำรวจ ก็คือ การที่ข้อมูลที่ได้มาจากกลุ่มตัวอย่างในการสำรวจนั้น สามารถใช้เป็นตัวแทนของกลุ่มประชากรทั้งหมดในประเด็นหรือปัญหาวิจัยที่กำลังศึกษาได้ ซึ่งทำให้หนึ่งในวิธีการที่ใช้ในการบ่งชี้ความคลาดเคลื่อนของชุดข้อมูลที่ได้มาจากการสำรวจได้ดีที่สุด ก็คือ วิธีการตรวจสอบโดยเปรียบเทียบกับข้อมูลภายนอก (External Validity) ที่มาจากกลุ่มประชากรทั้งหมดในประเด็นหรือปัญหาวิจัยที่กำลังศึกษา (Kojien, Van Nieuwerburgh & Vestman, 2014) ซึ่งหนึ่งในชุดข้อมูลที่มีลักษณะดังกล่าว ก็คือ “ข้อมูลเพื่อการบริหาร”

ข้อมูลเพื่อการบริหาร คือ ข้อมูลที่ได้จากการดำเนินงานหรือในระบบการบริหารจัดการ เพื่อจุดประสงค์ในการลงทะเบียน การซื้อขายแลกเปลี่ยน ตลอดจนการเก็บประวัติ เพื่อเป็นการให้บริการแก่ประชาชนหรือลูกค้าที่มาใช้บริการ (Woolard, 2014) ที่ผ่านเราจึงมักจะเห็นเพียงข้อมูลเพื่อการบริหารที่เผยแพร่โดยหน่วยงานภาครัฐเท่านั้น เนื่องจากข้อมูล

¹¹ การดำเนินการสำรวจจะเป็นการทยอยสำรวจเป็นประจำทุกเดือน แต่ทำการรวบรวมและเผยแพร่เป็นรายไตรมาส (4 ครั้งต่อปี) ขนาดการสุ่มกลุ่มตัวอย่างในแต่ละไตรมาสนั้น โดยเฉลี่ยแล้วมากกว่า 2 แสนคนจากทั่วประเทศ

¹² นอกจากนี้ ปัญหาในการนำข้อมูลแบบ Panel ซึ่งสร้างจากชุดข้อมูล LFS (เรียกอีกอย่างหนึ่งว่า Synthetic Panel Data) มาใช้ (Peracchi and Welch, 1995) และ ปัญหาทางเทคนิคที่เกี่ยวข้องกับการแบ่งกลุ่มตัวอย่าง เช่น ปัญหาความคลาดเคลื่อนอันเกิดจาก Outgoing Rotation Group หรือ กลุ่มตัวอย่างที่มีการสำรวจซ้ำ (Krueger, Mas and Niu, 2017; Ahn and Hamilton, 2020) ก็เป็นหนึ่งในหลายปัญหาที่นักวิจัยพบจากการใช้ชุดข้อมูลจากการสำรวจ

เพื่อการบริหารจากหน่วยงานภาคเอกชนมักจะ ไม่ถูกเผยแพร่ ด้วยเหตุผลในเรื่องข้อกฎหมายหรือการเก็บความลับทางการค้า ข้อมูลเพื่อการบริหารมักจะมีหน่วยการเก็บข้อมูลที่ละเอียดมาก เช่น ข้อมูลรายบุคคล ซึ่งส่งผลให้ข้อมูลเพื่อการบริหาร มักจะมีขนาดที่ใหญ่และครอบคลุมประชากรทั้งหมด (universal population coverage) หรือ $N = All$ นอกจากนี้ ด้วยการที่ข้อมูลเพื่อการบริหารมักครอบคลุมข้อมูลที่มีความอ่อนไหวและไม่สามารถเก็บข้อมูลด้วยวิธีการสำรวจได้ เช่น ข้อมูล การเกิด ข้อมูลการตาย ข้อมูลการเสียดาย ข้อมูลสุขภาพและการเจ็บป่วย เป็นต้น ด้วยเหตุนี้ การนำข้อมูลเพื่อการบริหารมา ใช้ในงานวิจัย จึงช่วยให้นักวิจัยสามารถทลายข้อจำกัดทางด้านรายละเอียดข้อมูล ลดความจำเป็นในการตั้งข้อสมมติฐานเพื่อ การทดสอบคุณสมบัติทางสถิติ และ เพิ่มโอกาสให้ผู้วิจัยสามารถประยุกต์ใช้การออกแบบงานวิจัย ในรูปแบบการวิจัยกึ่ง ทดลอง (Quasi-experimental research design) ได้มากยิ่งขึ้น (Einav and Levin, 2014)

ตรงกันข้ามกับข้อมูลที่ได้จากการสำรวจขนาดใหญ่ อย่างเช่น Current Population Survey (CPS) ในสหรัฐอเมริกา หรือ LFS ที่มีการสำรวจในหลายๆประเทศที่มีแนวโน้มที่ลดลงในการนำมาประยุกต์ใช้ในงานวิจัย การประยุกต์ใช้ข้อมูลเพื่อ การบริหารโดยเฉพาะอย่างยิ่งในงานวิจัยทางเศรษฐศาสตร์เริ่มเป็นที่แพร่หลายมากขึ้นในสองทศวรรษที่ผ่านมา¹³ เนื่องจาก ข้อได้เปรียบของขนาดของกลุ่มตัวอย่าง¹⁴ และมักไม่ค่อยพบปัญหาการขาดหายไปของข้อมูล อย่างไรก็ตาม ถึงแม้ว่าการใช้ ข้อมูลเพื่อการบริหารในงานวิจัยจะเริ่มเป็นที่แพร่หลายมากขึ้นในปัจจุบัน แต่งานวิจัยเหล่านี้ก็ยังคงจำกัดอยู่เฉพาะในบาง ปัญหาวิจัยที่เกี่ยวข้องกับข้อมูลเพื่อการบริหารเหล่านั้น และ/หรือ ประเทศที่หน่วยงานภาครัฐสนับสนุนการนำข้อมูลเพื่อการบริหาร ไปใช้ในการวิจัย ซึ่งเกือบทั้งหมดเป็นประเทศที่พัฒนาแล้ว อย่างเช่น สวีเดน (เช่น Besley, Folke, Persson & Rickne, 2017) เดนมาร์ก (เช่น Hummels, Jørgensen, Munch, & Xiang, 2014; Kleven, Landais, Saez, & Schultz, 2014) ฟินแลนด์ (Böckerman, Haapanen & Jepsen, 2018) และสหรัฐอเมริกา (Sullivan & Von Wachter, 2009; Chetty, Hendren, Kline, & Saez, 2014) สำหรับงานวิจัยในประเทศไทยที่มีการประยุกต์ใช้ข้อมูลเพื่อการบริหารนั้นพบว่ายังไม่เป็นที่แพร่หลายนัก เนื่องจากข้อจำกัดในการเข้าถึงและขอใช้ข้อมูลจากหน่วยงานเจ้าของข้อมูลซึ่งมักจะเป็นหน่วยงานภาครัฐ อย่างไรก็ตาม งานวิจัยที่น่าสนใจที่มีการใช้ข้อมูลเพื่อการบริหาร เช่น Apaitan & Wibulpolprasert (2018) ซึ่งใช้ข้อมูลเพื่อการบริหารการใช้ไฟฟ้ารายครัวเรือนจากการไฟฟ้าส่วนภูมิภาค Attavanich, Chantarat & Sa-ngimmet (2018) ซึ่งใช้ข้อมูลทะเบียนเกษตรกร และ Wasi, Paweenawat, Ayudhya, Treeratpituk & Nittayo (2019) ซึ่งใช้ข้อมูลที่สุ่มจากข้อมูลผู้ประกันตนรายบุคคลจาก สำนักงานประกันสังคม ซึ่งทั้งหมดก็มุ่งที่จะนำเสนอข้อเท็จจริงอันโดดเด่น (stylized facts) ในประเด็นปัญหาที่เกี่ยวข้องกับ ข้อมูล เท่านั้น

4.1 การเป็นตัวแทนประชากร (Representativeness) ของข้อมูล LFS

เนื่องจากข้อมูลผู้ประกันตนรายบุคคลจากสำนักงานประกันสังคม จะครอบคลุมกลุ่มประชากรที่เป็นแรงงานใน ภาคเอกชน เฉพาะที่เป็นการทำงานที่เป็นทางการ (formal employment) เท่านั้น แต่ข้อมูลจากชุดข้อมูล LFS จะมีความหมาย ครอบคลุม ทั้งการทำงานที่เป็นทางการ และ การทำงานที่ไม่เป็นทางการ (informal employment) ซึ่งทำให้เราควรจะ

¹³ Chetty (2012) ทำการรวบรวมจำนวนผลงานเชิงประจักษ์ (empirical research) ที่มีการตีพิมพ์ใน American Economic Review (ARE) Journal of Political Economy (JPE) Quarterly Journal of Economics (QJE) และ Econometrica (ECMA) พบว่าในปี ค.ศ. 2010 จำนวนผลงานเชิงประจักษ์ ที่ใช้ข้อมูลเพื่อการบริหารที่ตีพิมพ์ใน AER JPE และ QJE มีสัดส่วนอยู่ระหว่าง 60-75% ของงานเชิงประจักษ์ทั้งหมดที่ตีพิมพ์ในวารสารเหล่านี้ มี เพียง QJE เท่านั้น ที่จำนวนผลงานเชิงประจักษ์ที่ใช้ข้อมูลเพื่อการบริหารน้อยกว่า 40% ในขณะที่จำนวนผลงานเชิงประจักษ์ที่ใช้ข้อมูลจากการ สํารวจเช่น CPS (ไม่นับรวมงานที่ใช้ข้อมูลจากการสำรวจที่ออกแบบโดยเฉพาะสำหรับงานนั้นๆ) มีสัดส่วนอยู่ระหว่าง 10-25% เท่านั้น

¹⁴ ในสหรัฐอเมริกานั้น ขนาดของข้อมูลเพื่อการบริหารบางประเภทมีขนาดกว่า 2,000 เท่าของขนาดของข้อมูล CPS (Chetty, 2012)

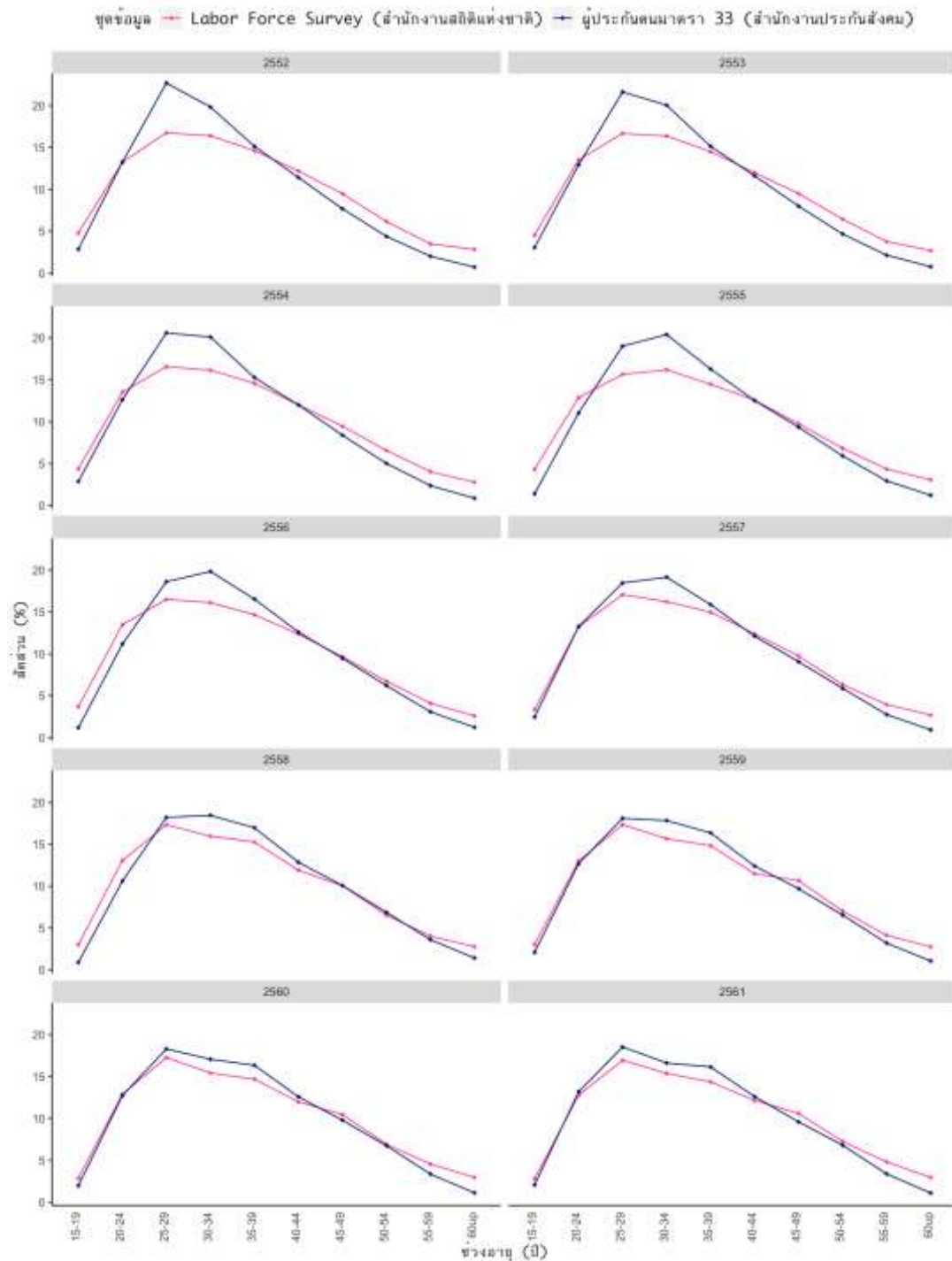
คาดหมายว่า “ค่าประมาณ” จำนวนแรงงานในภาคเอกชนที่ประมาณจากชุดข้อมูล LFS จะมีจำนวนที่มากกว่า “จำนวนแรงงานจริง” ที่ได้จากข้อมูลของสำนักงานประกันสังคม

รูปที่ 8 เปรียบเทียบสัดส่วน “ลูกจ้างเอกชน” ตามช่วงอายุระหว่างชุดข้อมูล LFS และชุดข้อมูลสถิติงานประกันสังคมซึ่งเป็นผลผลิตที่ได้โดยตรงจากสำนักงานประกันสังคม ที่ถือได้ว่าเป็นข้อมูลที่สมบูรณ์ที่สุดหากเรากำลังพิจารณาถึงลูกจ้างในภาคเอกชน การกระจายตัวที่คล้ายคลึงกันเป็นอย่างมากของแรงงานในกลุ่มอายุต่างๆที่ทำงานอยู่ในภาคเอกชน (P-value ของค่า Chi-squared ของทั้งสองปีอยู่ระหว่าง 0.91~0.99) ผลดังกล่าวไม่เป็นที่น่าแปลกใจแต่อย่างใด เนื่องจากชุดข้อมูล LFS ใช้แผนการเลือกตัวอย่าง (รวมถึงค่าถ่วงน้ำหนัก) ที่เป็นไปตามโครงสร้างจำนวนประชากรในแต่ละช่วงอายุ ที่อาศัยอยู่ในแต่ละภาค และเขตการปกครองเป็นสำคัญ

อย่างไรก็ตาม เมื่อเปรียบเทียบจำนวนของแรงงานในภาคเอกชนโดยใช้ช่วงรายได้แล้ว เราจะเริ่มเห็นความผิดปกติของ “ค่าประมาณ” จำนวนแรงงานในภาคเอกชนที่คำนวณจากชุดข้อมูล LFS **ตารางที่ 5** แสดงผลต่างจำนวนแรงงานในภาคเอกชนที่คำนวณโดยใช้ “ค่าประมาณ” ของชุดข้อมูล LFS ลบด้วย “จำนวนแรงงานจริง” ที่ได้จากข้อมูลของสำนักงานประกันสังคม ซึ่งโดยนิยามแล้ว ผลต่างจำนวนแรงงานในภาคเอกชนที่ประมาณการจากชุดข้อมูล LFS ควรจะมีจำนวนที่มากกว่าหรือเท่ากับข้อมูล “จำนวนแรงงานจริง” ที่ได้จากข้อมูลผู้ประกันตนรายบุคคลของสำนักงานประกันสังคม อย่างไรก็ตาม ผลต่างที่ติดลบแสดงถึงปัญหาที่ชัดเจนของข้อสมมติฐาน (หรือความเชื่อ) ที่ว่า “ข้อมูลที่มาจากการสำรวจขนาดใหญ่สามารถใช้เป็นตัวแทนของประชากรได้” โดยตัวอย่างหนึ่งที่ชัดเจนก็คือ จำนวนลูกจ้างเอกชนในช่วงรายได้มากกว่า 13,000 บาทต่อเดือน ยกตัวอย่างเช่น ปี 2561 นั้น การประมาณจากข้อมูล microdata ในไตรมาส 4 ของ LFS แสดงให้เห็นว่าจำนวนลูกจ้างทั้งหมดในภาคเอกชน (ทั้งแบบเป็นทางการและไม่เป็นทางการ) จะมีจำนวน 13,891,627 คน โดยมีจำนวนลูกจ้างที่ได้รายได้มากกว่า 13,000 บาทต่อเดือนเท่ากับ 4,030,336 คน แต่ข้อมูลจากสำนักงานประกันสังคมแสดงให้เห็นว่า ณ สิ้นปี 2561 จำนวนลูกจ้างทั้งหมดแบบเป็นทางการในภาคเอกชนมีจำนวนทั้งสิ้น 11,599,338 คน และมีจำนวนลูกจ้างเอกชนในช่วงรายได้มากกว่า 13,000 บาทต่อเดือนทั้งสิ้น 4,655,607 คน ซึ่งความแตกต่างกว่า 625,271 คน นี้สูงมากจนไม่อาจกล่าวได้ว่าเป็นความคลาดเคลื่อนที่เป็นปรกติในทางสถิติ¹⁵ ซึ่งแสดงให้เห็นว่า หากเราใช้ค่าประมาณที่ได้จากชุดข้อมูล LFS มาใช้ในงานวิจัยในประเด็นที่เกี่ยวข้องกับรายได้จากการทำงาน เช่น ผลตอบแทนทางการศึกษา ค่าประมาณที่คลาดเคลื่อนดังกล่าวก็อาจจะส่งผลให้การประมาณค่าผลตอบแทนทางการศึกษา (โดยเฉพาะปริญญาตรี) มีความคลาดเคลื่อน เนื่องจากลูกจ้างในส่วนของกลุ่มลูกจ้างเอกชนที่มีรายได้ในช่วงรายได้สูงกว่า 13,000 บาทต่อเดือนในปัจจุบันน่าจะเป็นช่วงรายได้ที่แรงงานจะต้องมีการศึกษาในระดับปริญญาตรีขึ้นไป

นอกจากข้อมูลที่เกี่ยวข้องกับการทำงานแล้ว การสำรวจภาวะการทำงานของประชากรยังเป็นการสำรวจที่ให้ความสำคัญกับรายละเอียดทางด้านการศึกษา ซึ่งเป็นองค์ประกอบที่สำคัญต่อการวิเคราะห์ด้านแรงงาน ด้วยเหตุนี้ เราจึงสามารถทำการตรวจสอบความถูกต้องของชุดข้อมูล LFS ในด้านการศึกษาได้อีกทางหนึ่ง งานวิจัยนี้เลือกใช้สถิติจำนวนผู้จบการศึกษาในระดับชั้นต่างๆ จากกระทรวงศึกษาธิการ กระทรวงแรงงาน และสำนักงานคณะกรรมการการอุดมศึกษา เนื่องจากเป็นหน่วยงานที่เป็นผู้จัดเก็บข้อมูลและเผยแพร่สถิติเหล่านี้ ทั้งโดยตรงและโดยอ้อม

¹⁵ ถึงแม้ว่าข้อมูล microdata จะไม่มีข้อมูลค่าถ่วงน้ำหนักที่สามารถนำไปคำนวณช่วงความเชื่อมั่นของค่าประมาณได้ อย่างไรก็ตาม หากพิจารณารายละเอียดช่วงความเชื่อมั่น (95%) ที่แสดงในตารางคุณลักษณะของประชากรและตัวอย่าง ในรายงานการสำรวจภาวะการทำงานของประชากรรายไตรมาส ที่เผยแพร่เป็นประจำของสำนักงานสถิติแห่งชาติแล้ว เราสามารถใช้ช่วงความเชื่อมั่นดังกล่าวมาคาดการณ์ช่วงความเชื่อมั่นของค่าประมาณแบบหยาบได้



รูปที่ 8: สัดส่วนลูกจ้างเอกชน ตามช่วงอายุ ของข้อมูล LFS และข้อมูลผู้ประกันตน

หมายเหตุ:

1. จำนวนลูกจ้างเอกชนจากชุดข้อมูล LFS จำนวน โดยใช้ microdata ของไตรมาสที่ 4 ในแต่ละปี จากสำนักงานสถิติแห่งชาติ โดยกำหนดให้อายุมากกว่าเท่ากับ 15 ปี ($AGE \geq 15$) สถานะการทำงานลูกจ้างเอกชน ($WKSTAT = 6$) และลักษณะการทำงานอยู่ในกลุ่มที่มีงานทำ ($RE_WK = \{1,2,3\}$)
2. ข้อมูลจำนวนผู้ประกันตนมาตรา 33 จากสำนักงานประกันสังคม
3. P-value ของค่า Chi-squared ของทั้งสิบปีอยู่ระหว่าง 0.91~0.99

ตารางที่ 5: ผลต่างจำนวนแรงงานภาคเอกชนระหว่างข้อมูล LFS และข้อมูลจากสำนักงานประกันสังคม แยกตามระดับรายได้ ตั้งแต่ปี 2552-2561 (หน่วย: คน)

ช่วงรายได้	ปี พ.ศ.									
	2552	2553	2554	2555	2556	2557	2558	2559	2560	2561
1-1650	550,235	254,159	69,014	84,124	-1,386	-39,880	-47,656	-59,249	-69,943	-84,361
1651-2000	262,501	223,598	162,639	152,762	36,837	30,211	49,080	42,731	41,287	50,748
2001-3000	1,059,909	692,898	516,735	382,700	190,259	155,660	163,884	191,515	175,449	195,789
3001-4000	1,070,192	770,180	671,446	629,880	201,412	181,807	162,493	196,338	147,979	139,126
4001-5000	1,169,819	1,119,706	1,123,454	887,026	460,834	288,888	260,237	212,516	253,604	190,728
5001-6000	793,053	938,367	992,227	1,125,788	517,384	453,355	388,250	337,168	348,240	349,769
6001-7000	-26,718	48,243	120,656	484,462	-118,828	102,139	-70,349	-55,416	-60,685	-25,111
7001-8000	353,432	320,336	528,269	1,013,349	1,222,067	1,640,152	1,544,336	1,481,339	792,819	437,338
8001-9000	-59,189	20,365	100,811	-293,480	193,432	193,908	323,668	330,470	1,087,134	1,109,850
9001-10000	84,199	93,416	91,472	2,451	93,599	326,984	323,239	368,985	41,847	30,954
10001-11000	-137,388	-162,348	-171,028	-222,604	-194,061	-223,754	-160,121	-109,626	-6,995	-104,134
11001-12000	45,036	56,212	71,567	132,347	313,282	491,960	561,235	590,338	584,450	558,683
12001-13000	-48,946	-60,030	-56,122	-98,650	-72,878	-29,649	-17,047	12,671	49,352	68,181
13001-14000	-75,357	-111,152	-105,293	-119,988	-128,627	-100,716	-197,369	-237,143	-240,918	-206,098
14001 ขึ้นไป	-306,795	-326,177	-321,147	-251,991	-188,870	205,609	185,039	31,111	-288,682	-419,173

หมายเหตุ:

- จำนวนลูกจ้างเอกชนจากชุดข้อมูล LFS จำนวน โดยใช้ microdata ของไตรมาสที่ 4 ในแต่ละปี จากสำนักงานสถิติแห่งชาติ โดยกำหนดค่าอายุมากกว่าเท่ากับ 15 ปี ($AGE \geq 15$) สถานะการทำงานลูกจ้างเอกชน ($WKSTAT = 6$) และลักษณะการทำงานอยู่ในกลุ่มที่มีงานทำ ($RE_WK = \{1,2,3\}$)
- ข้อมูลจำนวนผู้ประกันตนมาตรา 33 จากสำนักงานประกันสังคม
- ผลต่างจำนวนแรงงานในภาคเอกชน เท่ากับ จำนวนลูกจ้างเอกชนจากชุดข้อมูล LFS ลบด้วย จำนวนแรงงานจากข้อมูลจำนวนผู้ประกันตนจากสำนักงานประกันสังคม

ตารางที่ 6: จำนวนผู้จบการศึกษา แยกตามระดับชั้น ตั้งแต่ปี 2543-2561 (หน่วย: คน)

ปีที่จบการศึกษา พ.ศ.	ม.3	ม.6	ปวช. (ข้อมูล MOE)	ปวช. (ข้อมูล MOL)	ปวส. (ข้อมูล MOE)	ปวส. (ข้อมูล MOL)	ปริญญาตรี
ลำดับศตมภ์	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2543	714,610**	289,536**	201,829*	160,766**	212,593*	133,950**	154,388**
2544	756,726**	344,362**	175,941*	172,046**	190,499*	190,174**	144,183**
2545	714,426**	351,854**	157,604*	142,676**	195,017*	157,627**	154,113**
2546	716,738*	336,665*	149,621*	135,886**	183,770*	179,613**	152,679**
2547	748,149*	345,525*	147,960*	158,359**	170,680*	187,537**	204,495**
2548	787,470*	328,462*	155,818*	138,321**	158,692*	145,124**	195,798**
2549	848,404*	319,049*	156,148*	152,728**	171,688*	129,417**	195,077**
2550	871,937*	320,351*	181,716*	169,182**	182,387*	146,425**	206,892**
2551	864,503*	340,760*	182,899*	199,226**	173,485*	174,404**	252,122**
2552	862,806*	355,304*	200,156*	219,642**	164,184*	115,245**	263,489**
2553	866,079*	376,133*	207,368*	231,861**	165,512*	124,612**	272,281**
2554	899,506*	388,836*	211,821*	231,445**	173,035*	152,987**	286,632**
2555	898,608*	408,541*	205,540*	224,360**	170,857*	182,602**	280,253***
2556	846,135*	422,749*	203,909*	223,263**	147,989*	160,598**	283,153***
2557	778,767*	449,110*	211,766*	231,179**	147,921*	160,141**	308,693***
2558	772,386*	472,915*	211,701*	197,186**	142,582*	189,917**	316,751***
2559	769,329*	471,205*	192,789*	144,720**	152,917*	122,494**	357,667***
2560	739,735*	416,080*	183,350*	80,875**	159,423*	91,316**	319,801***
2561	739,698*	402,865*	194,683*	124,732**	162,861*	125,499**	330,393***
จำนวนรวม	15,926,837	7,540,363	3,723,818	3,486,592	3,391,255	2,973,069	4,953,800

หมายเหตุ:

- * ใช้ข้อมูลจากรายงานสถิติการศึกษาลบข้อ ซึ่งจัดทำเป็นประจำทุกปีจัดทำโดยกระทรวงศึกษาธิการ (MOE)
- ** ใช้ข้อมูลจากรายงานผู้สำเร็จการศึกษาและประมาณการผู้เข้าสู่ตลาดแรงงาน ซึ่งจัดทำเป็นประจำทุกปีโดยกระทรวงแรงงาน (MOL)
- *** ใช้ข้อมูลจากรายงานจำนวนผู้สำเร็จการศึกษา เว็บไซต์สารสนเทศอุดมศึกษา โดยสำนักงานคณะกรรมการการอุดมศึกษา
- จำนวนผู้จบการศึกษาระดับชั้น ปวช. ในลำดับศตมภ์ที่ (3) จำนวนโดยใช้จำนวนนักศึกษาที่กำลังศึกษาอยู่ ณ ตอนต้นปีการศึกษาที่ผ่านมา (จำนวนรวมทั้งสิ้น 3,952,960 คน) และคูณด้วยค่าคงที่ 0.89 ซึ่งคำนวณจากจำนวนนักศึกษาในระดับ ปวช. ที่ออกกลางคันระหว่างปี ซึ่งมีค่าเฉลี่ยเท่ากับ 0.11 (ข้อมูลจำนวนนักศึกษาที่ออกกลางคันระหว่างปี มีรายงานเฉพาะในปีการศึกษา 2557 - 2560 เท่านั้น)

ตารางที่ 6 แสดงจำนวนผู้จบการศึกษาในแต่ละปีแยกตามระดับชั้น ตั้งแต่ปี พ.ศ. 2543-2561 โดยใช้ข้อมูลจากทั้งสามหน่วยงานดังที่ได้กล่าวถึงในส่วนที่ผ่านมา เนื่องจากความไม่สมบูรณ์ของข้อมูลจากแต่ละแหล่งข้อมูล ทั้งนี้ในระดับ ปวช. และ ปวส. นั้น งานวิจัยนี้ทำการเลือกแสดงข้อมูลจากทั้งกระทรวงศึกษาธิการและกระทรวงแรงงาน เนื่องจากพบความแตกต่างที่น่าจะมีนัยสำคัญ (โดยเฉพาะระดับ ปวส.) ที่อาจส่งผลกระทบต่อเปรียบเทียบในชุดข้อมูลจากทั้งสองหน่วยงาน ผลรวมของผู้จบการศึกษาจาก**ตารางที่ 6** ถูกนำไปคำนวณกรอบจำนวนผู้จบการศึกษาในระดับมัธยมศึกษาตอนต้น (Lower Secondary Level) ระดับมัธยมศึกษาตอนปลาย (Upper Secondary Level) และระดับประกาศนียบัตรวิชาชีพชั้นสูงขึ้นไป

เนื่องจากผู้จบการศึกษาในระดับชั้น ม.6 และ ปวช. สามารถเลือกทำการศึกษาต่อได้ทั้งในระดับ ปวส. และระดับปริญญาตรี นอกจากนี้ ผู้จบการศึกษาในระดับ ปวส. ก็ยังสามารถทำการศึกษาต่อในระดับปริญญาตรี ด้วยเหตุนี้ หากจุดประสงค์สูงสุด คือ การคำนวณหาจำนวน “ผู้ที่มีระดับการศึกษาสูงสุดในระดับปริญญาตรี ในปี 25XX” การนำจำนวนผู้จบการศึกษาในระดับชั้น ม.6 ในปีการศึกษา $t-1$ และลบด้วยจำนวนนักศึกษาปริญญาตรีชั้นปีที่ 1 ในปีการศึกษา t จึงไม่ถูกต้องนัก งานวิจัยนี้ จึงทำการเปรียบเทียบจำนวนรวมทั้งหมดของ “ผู้ที่มีระดับการศึกษาสูงสุด” ในระดับชั้นมัธยมศึกษาตอนต้น ระดับชั้นมัธยมศึกษาตอนปลาย (ซึ่งเป็นผลรวมของผู้ที่มีระดับการศึกษาสูงสุด ทั้งที่อยู่สายสามัญ สายอาชีวศึกษา และสายวิชาการศึกษา) และระดับที่สูงกว่ามัธยมศึกษาตอนปลายทั้งหมด (ปวส. และสูงกว่า ปวส.) ที่จบการศึกษาระหว่างปี พ.ศ. 2543 ถึง 2561 แทนการเปรียบเทียบจำนวนผู้จบการศึกษาระดับในแต่ละปี ในแต่ละปี นอกจากนี้ วิธีการนี้ยังช่วยการลดปัญหาการนับซ้ำ เนื่องจากผู้จบการศึกษาในแต่ละระดับชั้นในแต่ละปี อาจไม่ได้ทำการศึกษาต่อโดยทันทีและอาจศึกษาต่อในระดับชั้นที่สูงขึ้นไปในภายหลัง โดยวิธีการดังกล่าวก็จะถูกนำไปใช้กับ microdata ของชุดข้อมูล LFS ด้วยเช่นกัน เพื่อทำการคำนวณตัวเลขค่าประมาณ “จำนวนประชากร จำแนกตามระดับการศึกษาสูงสุด” ดังแสดงใน **ตารางที่ 7**

ตารางที่ 7: กรอบจำนวนประชากร จำแนกตามระดับการศึกษาสูงสุดที่สำเร็จ ระหว่างปี 2543 ถึง 2561

	ม.3 (ข้อมูล MOE)	ม.6 + ปวช. (ข้อมูล MOE)	ม.6 + ปวช. (ข้อมูล MOL)	ปวส. ขึ้นไป (ข้อมูล MOE)	ปวส. ขึ้นไป (ข้อมูล MOL)
วิธีการคำนวณ	(1) - (2) - (3)	(2) + (3) - (5) - (7)	(2) + (4) - (6) - (7)	(5) + (7)	(6) + (7)
จำนวน (คน)	4,662,655	2,919,126	3,100,086	8,345,055	7,926,869
ข้อมูล LFS (คน)	7,541,948	6,237,077		6,143,537	

หมายเหตุ:

- ข้อมูล LFS จำนวน โดยใช้ microdata ของไตรมาส 4 ปี พ.ศ. 2561 โดยกำหนดให้ตัวแปรอายุ (AGE) อยู่ระหว่าง 16-41 ปี ($41 \geq AGE \geq 16$) ซึ่งมาจากสมมติฐานอายุขั้นต่ำในการจบการศึกษาในแต่ละระดับชั้น ตามเกณฑ์อายุปกติ (เข้าเรียนชั้น ป.1 เมื่ออายุได้ 6 ปีบริบูรณ์หรืออย่างช้าปีที่ 7) คือ ม.3 เท่ากับ 15 ปี ม.6 และ ปวช. เท่ากับ 18 ปี ปวส. เท่ากับ 20 ปี และปริญญาตรี เท่ากับ 22 ปี ซึ่งเมื่อคำนวณจากปี พ.ศ. 2561 ย้อนกลับไปที่ พ.ศ. 2543 (19 ปี) จะได้กรอบอายุที่น้อยที่สุดในปี พ.ศ. 2561 ของผู้ที่จบปริญญาตรีในปี พ.ศ. 2543 เท่ากับ 41 ปี และกรอบอายุที่น้อยที่สุดในปี พ.ศ. 2561 ของผู้ที่จบ ม.3 ในปี พ.ศ. 2561 เท่ากับ 16 ปี
- การกำหนดระดับการศึกษาสูงสุด (ตัวแปร RE_ED) จาก microdata ของชุดข้อมูล LFS สำหรับการคำนวณจำนวน “ผู้ที่มีระดับการศึกษาสูงสุด” ในแต่ละระดับ สามารถกำหนดด้วย Code ดังต่อไปนี้
 - ม.3 จะกำหนดให้ RE_ED เท่ากับ 4
 - ม.6 + ปวช. จะกำหนดให้ RE_ED อยู่ในช่วง 5 - 7
 - ปวส. ขึ้นไป จะกำหนดให้ RE_ED อยู่ในช่วง 8 - 15
- วิธีการคำนวณในตารางนี้ อ้างอิงข้อมูล “จำนวนรวม” ซึ่งปรากฏในแถวสุดท้ายของแต่ละลำดับสมมติในตารางที่ 4.2 เช่น (1) หมายถึง ลำดับสมมติที่ (1) ซึ่งเป็นสมมติของจำนวนผู้จบการศึกษาระดับชั้น ม.3 เป็นต้น

จำนวนจำนวนประชากร จำแนกตามระดับการศึกษาสูงสุดที่สำเร็จในแต่ละระดับที่เข้าสู่ตลาดแรงงาน ซึ่งคำนวณโดยใช้ microdata จากชุดข้อมูล LFS แสดงให้เห็นถึงความคลาดเคลื่อนเป็นอย่างมาก เมื่อเปรียบเทียบกับกรอบจำนวนประชากร จำแนกตามระดับการศึกษาสูงสุดที่สำเร็จ ที่คำนวณโดยใช้สถิติจำนวนผู้จบการศึกษาจากหน่วยงานต่างๆของรัฐที่เป็นผู้จัดเก็บข้อมูล ซึ่งความคลาดเคลื่อนมีขนาดใหญ่เป็นอย่างมากจนไม่อาจกล่าวได้ว่าเป็นความคลาดเคลื่อนที่เป็นปกติในทางสถิติ และเป็นอีกข้อบ่งชี้ของปัญหาการเป็นตัวแทนประชากรของชุดข้อมูล LFS นอกจากนี้ งานวิจัยนี้ยังได้ใช้ (1) ข้อมูลจำนวนผู้จบการศึกษาที่รวบรวมจากสำนักงานคณะกรรมการอุดมศึกษา กรมการจัดหางาน และแผนพัฒนาการศึกษา

ระดับอุดมศึกษา เพื่อตรวจสอบกรอบจำนวนแรงงานที่จบการศึกษาในระดับปริญญาตรีขึ้นไป ตั้งแต่ปี พ.ศ. 2520-2561 และ (2) ข้อมูลกำลังคนภาครัฐจากสำนักงานคณะกรรมการข้าราชการพลเรือน และข้อมูลลูกจ้างจากหน่วยงานรัฐวิสาหกิจต่างๆ (ทั้งจากข้อมูลที่มีการเผยแพร่แก่สาธารณะ เช่น รายงานประจำปี เว็บไซต์ของหน่วยงาน รวมถึงการโทรศัพท์เพื่อสอบถาม โดยตรงถึงหน่วยงานที่ไม่มีการเผยแพร่ข้อมูล) เพื่อตรวจสอบจำนวนลูกจ้างรัฐบาลและรัฐวิสาหกิจ สำหรับปี พ.ศ. 2561 ผลการเปรียบเทียบสถิติ พบความแตกต่างเป็นอย่างมากจนไม่อาจกล่าวได้ว่าเป็นความคลาดเคลื่อนที่เป็นปรกติในทางสถิติ โดยการประมาณจำนวนแรงงานที่จบการศึกษาในระดับปริญญาตรีขึ้นไป และ จำนวนจำนวนลูกจ้างรัฐบาลและรัฐวิสาหกิจ จากข้อมูล microdata ของ LFS ในไตรมาส 4 ของปี 2561 มีจำนวนทั้งสิ้น 7,492,018 และ 3,486,765 คน ตามลำดับ เมื่อเทียบกับจำนวน 6,109,947 และ 2,746,537 คน เมื่อใช้ข้อมูลสถิติที่มีการเผยแพร่

ถึงแม้ว่าผลการตรวจสอบจำนวนลูกจ้างในภาคเอกชน จำนวนของผู้จบการศึกษาสูงสุดในระดับต่างๆ ตลอดจนจำนวนลูกจ้างรัฐบาลและรัฐวิสาหกิจ บ่งชี้อย่างชัดเจนถึงปัญหาในการใช้ “ค่าประมาณ” และปัญหาการเป็นตัวแทนประชากรของชุดข้อมูล LFS อย่างไรก็ตาม การอภิปรายในส่วนนี้มิได้มีเจตนาในการนำเสนอว่าข้อมูล LFS ผิดหรือไม่มีประโยชน์แต่อย่างใด ในทางตรงกันข้าม ผู้วิจัยเห็นว่าข้อมูล LFS โดยเฉพาะอย่างยิ่งข้อมูลดิบหรือข้อมูล microdata ของชุดข้อมูลดังกล่าว นั้น มีประโยชน์เป็นอย่างมากต่อการนำไปสร้างนโยบาย หรือ การนำมาประยุกต์ใช้ในการวิจัย เนื่องจากชุดข้อมูล LFS นั้นประกอบไปด้วยข้อมูลหลายส่วนที่ช่วยให้การวิเคราะห์สามารถทำได้อย่างสมบูรณ์มากยิ่งขึ้น ดังที่ได้อภิปรายในส่วนก่อนหน้า อย่างไรก็ตาม ประเด็นปัญหาที่สำคัญของการอภิปรายในส่วนนี้ คือ การที่นักวิจัยที่นำข้อมูล LFS ไปใช้ มักจะขาดความระมัดระวังในการนำข้อมูลที่เป็น “ผลผลิต” (derivation) ที่คำนวณจากข้อมูลที่ได้ที่ได้ออกการสำรวจ (survey-based) โดยเฉพาะในระดับรวม (Aggregate level) ไปใช้ โดยหลงลืมไปว่าข้อมูลที่เป็นผลผลิตในระดับรวมดังกล่าว เป็นเพียง “ค่าประมาณ” ที่ได้มาโดยอ้อมจากการคำนวณขึ้นมาโดยใช้ค่าถ่วงน้ำหนัก ซึ่งค่าถ่วงน้ำหนักดังกล่าวมีความสัมพันธ์โดยตรงกับกรอบประชากรและวิธีการเก็บข้อมูล (ซึ่งสำหรับชุดข้อมูล LFS นั้น สำนักงานสถิติแห่งชาติใช้แผนการเลือกตัวอย่าง ที่เป็นไปตามโครงสร้างจำนวนประชากรในแต่ละช่วงอายุ ในแต่ละภาคและเขตการปกครองเป็นสำคัญ) ซึ่งหากในความเป็นจริงแล้ว ปัจจัยที่ส่งผลต่อการกระจายตัวของข้อมูลและประชากรในประเด็นที่ต้องการวิเคราะห์ ไม่ได้เป็นไปตามโครงสร้างจำนวนประชากรในแต่ละช่วงอายุ ในแต่ละภาค และในแต่ละเขตการปกครอง ตามแผนการเลือกตัวอย่างที่ได้ออกแบบไว้ ค่าถ่วงน้ำหนักเหล่านี้ก็จะทำให้ “ค่าประมาณ” ของระดับรวมที่ได้ นั้น ผิดเพี้ยนไปด้วย เช่นเดียวกัน

4.2 ข้อควรระวังจากปัญหาของการกระจุกตัวของกลุ่มตัวอย่างของข้อมูล ORE

ตารางที่ 8 แสดงสัดส่วนของกลุ่มตัวอย่างในข้อมูล ORE แยกตามสาขาวิชาในระดับปริญญาตรี โดยสาขาที่มีผู้จบมากที่สุดสามลำดับแรก คือ กลุ่มสาขาวิชาสังคมศาสตร์และบริหารธุรกิจ (โดยเฉลี่ย 50%) กลุ่มสาขาวิชาวิศวกรรมศาสตร์ (โดยเฉลี่ย 24%) และกลุ่มสาขาวิชาการบริการ โรงแรมและการท่องเที่ยว (โดยเฉลี่ย 6%) ซึ่งสัดส่วนดังกล่าวแตกต่างจากสัดส่วนของจำนวนผู้จบการศึกษาจริงในกลุ่มสาขาวิชาเหล่านี้ (34% 9% และ 3% ตามลำดับ)¹⁶ ซึ่งจากการรวบรวมจำนวนงานทั้งสามประเภท ที่ลงโฆษณาในเว็บไซต์ของผู้ให้บริการรายหนึ่งเพื่อตรวจสอบในเบื้องต้น งานวิจัยนี้พบว่าสัดส่วนจำนวนของงานแต่ละประเภทที่โฆษณานั้น มีสัดส่วนที่ใกล้เคียงกับสัดส่วนของข้อมูล ORE

¹⁶ คำนวณโดยใช้ข้อมูลจากสถิติการศึกษาปี 2561 ที่มา: <http://www.mis.moe.go.th/>

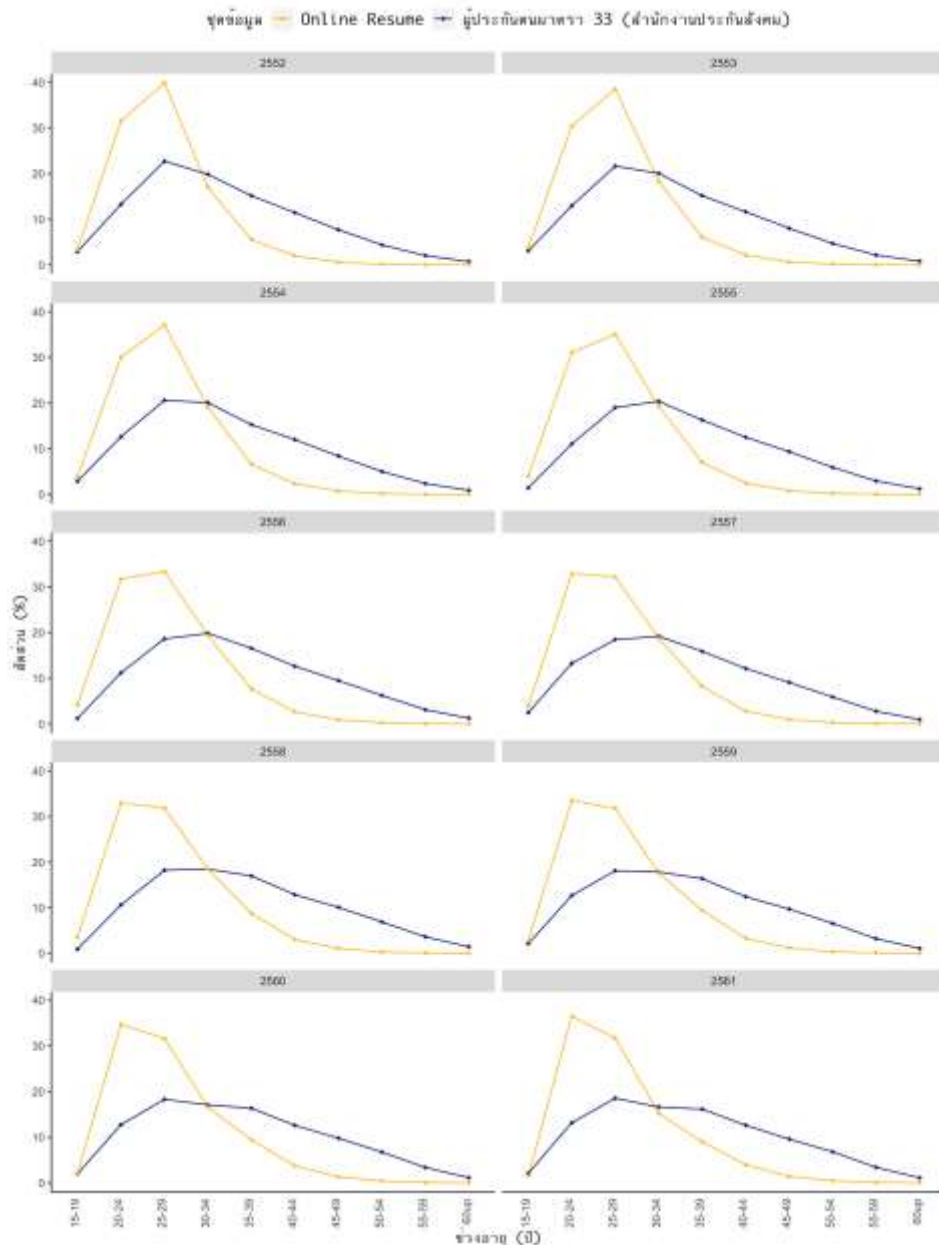
ตารางที่ 8: สัดส่วนกลุ่มตัวอย่างของข้อมูล ORE แยกตามสาขาวิชาในระดับปริญญาตรี (หน่วย: เปอร์เซนต์)

ปี พ.ศ. ที่ จบการศึกษา	EDU	HUM	ART	LAW	SOC	SCI	MED	ENG	AGR	SER
2542	3.35	3.52	0.99	1.81	62.71	2.33	2.30	18.61	2.18	2.18
2543	3.16	3.11	0.63	1.67	61.96	2.53	1.90	21.22	1.69	2.12
2544	3.35	3.77	0.69	1.90	60.72	3.16	1.83	19.43	2.63	2.53
2545	3.57	3.66	0.78	1.89	60.97	2.44	1.94	19.31	2.73	2.70
2546	3.42	3.53	0.67	2.33	59.94	2.32	2.25	19.02	3.18	3.33
2547	2.81	3.45	0.67	2.55	60.54	2.54	1.76	19.35	2.97	3.35
2548	2.46	3.96	0.65	2.87	58.79	2.21	1.70	20.09	3.39	3.87
2549	1.85	3.64	0.66	3.46	57.81	2.08	1.82	20.84	3.51	4.34
2550	1.36	3.88	0.51	3.54	56.88	2.22	1.53	22.43	3.03	4.62
2551	0.95	3.99	0.65	3.65	54.48	2.14	1.60	24.77	2.70	5.07
2552	1.18	4.35	0.64	3.89	51.49	2.15	1.55	27.01	2.54	5.20
2553	1.44	4.90	0.70	3.80	47.78	2.77	1.62	29.16	2.64	5.20
2554	1.28	4.76	0.65	3.87	46.67	3.12	1.85	29.42	2.98	5.40
2555	1.62	4.49	0.72	3.95	45.47	3.65	1.97	28.57	3.33	6.22
2556	2.34	5.34	0.79	3.88	43.17	4.83	2.25	27.08	3.83	6.48
2557	3.02	5.49	0.94	3.95	44.68	4.97	1.90	24.95	3.33	6.79
2558	4.84	6.29	0.87	3.70	42.28	5.87	1.95	24.04	3.58	6.57
2559	5.15	6.10	0.94	3.49	44.44	5.34	1.98	22.38	3.49	6.69
2560	4.20	5.86	0.94	2.87	47.84	5.77	2.39	20.29	3.47	6.38
2561	3.68	6.15	0.89	2.32	48.56	5.48	2.61	19.90	3.55	6.86

หมายเหตุ:

1. EDU หมายถึง สาขาวิชาศึกษาศาสตร์ HUM หมายถึง สาขาวิชามนุษยศาสตร์ ART หมายถึง สาขาวิชาศิลปกรรมศาสตร์ LAW หมายถึง สาขาวิชานิติศาสตร์ SOC หมายถึง สาขาวิชาสังคมศาสตร์และบริหารธุรกิจ SCI หมายถึง สาขาวิชาวิทยาศาสตร์กายภาพ MED หมายถึง สาขาวิชาการแพทย์และสาธารณสุข ENG หมายถึง สาขาวิชาวิศวกรรมศาสตร์ AGR หมายถึง สาขาวิชาเกษตรศาสตร์ และ SER หมายถึง สาขาวิชาบริการโรงแรมและการท่องเที่ยว

หนึ่งในคำอธิบายปรากฏการณ์ที่เกิดขึ้นนี้ คือ การที่องค์กรธุรกิจจะทำการโฆษณาเพื่อรับสมัครแรงงานนั้น องค์กรธุรกิจจะต้องเสียค่าใช้จ่าย (ซึ่งตรงข้ามกับผู้ใช้บริการที่เป็นแรงงานที่กำลังหางาน ซึ่งไม่จำเป็นต้องเสียค่าใช้จ่ายในการฝากประวัติหรือสมัครงานออนไลน์) ทำให้ความครอบคลุมประเภทของงานที่มักจะเข้ามาโฆษณาเพื่อรับสมัครงาน ถูกจำกัดอยู่กับเฉพาะงานบางประเภท อุตสาหกรรมบางประเภท ตลอดจน ขนาดขององค์กรธุรกิจที่ต้องสามารถแบกรับต้นทุนที่เกิดจากการโฆษณารับสมัครงานได้ (Cajner & Ratner, 2016) นอกจากนี้ ด้วยการทำงานบางประเภทอาจจะไม่มีความจำเป็นที่จะต้องประกาศรับสมัครออนไลน์ หรือแม้แต่ว่าจะได้รับการตอบรับที่ดีกว่า หากโฆษณาผ่านสื่อประชาสัมพันธ์ชนิดอื่น เช่น หนังสือพิมพ์ เว็บไซต์ของหน่วยงาน หรือแม้แต่ประกาศหน้าร้าน (Turrell et al., 2019) ดังนั้น จึงไม่น่าแปลกใจแต่อย่างใด หากสัดส่วนของแรงงานที่เข้ามาหางานในข้อมูล ORE จะมีจำนวนที่สูงมาก ในสายงานที่มีจำนวนคนจบการศึกษาไม่สูงในแต่ละปีหากแต่เป็นที่ต้องการในองค์กรภาคเอกชนเป็นอย่างมาก เช่น สาขาวิชาวิศวกรรมศาสตร์ แต่ในทางตรงกันข้าม



รูปที่ 9: การเป็นตัวแทนประชากรของข้อมูล ORE ตามช่วงอายุ

หมายเหตุ:

1. ข้อมูลจำนวนผู้ประกันตนมาตรา 33 จากเว็บไซต์สำนักงานประกันสังคม
2. ข้อมูลจำนวนลูกจ้างเอกชนจากชุดข้อมูล ORE โดยจะทำการใช้ข้อมูลเฉพาะผู้ที่มีประวัติการทำงานในเดือนธันวาคมของปีที่พิจารณาเท่านั้น
3. P-value ของค่า Chi-squared ของทั้งสิบปีอยู่ระหว่าง 9.2×10^{-7} - 3.04×10^{-5}

แรงงานในสายงานทางด้านวิทยาศาสตร์ที่เข้ามาหางานทางอินเทอร์เน็ตกลับมีสัดส่วนที่ต่ำมาก ยิ่งเมื่อพิจารณาถึงจำนวนคนจบการศึกษาในแต่ละปี ที่เป็นรองเพียงแก่สาขาวิชาทางด้านสังคมศาสตร์และบริหารธุรกิจเท่านั้น เนื่องจากกลุ่มสาขาวิทยาศาสตร์เป็นสาขาที่มักไม่เป็นที่ต้องการในองค์กรภาคเอกชน ดังนั้น การที่อุปสงค์แรงงาน (รวมถึงตำแหน่งงาน) ในการหาแรงงานทางอินเทอร์เน็ตถูกจำกัดและครอบงำโดยอุตสาหกรรมบางประเภทเช่นนี้ ทำให้การกระจายตัวของข้อมูล ORE ในงานวิจัยนี้ มีการกระจุกตัวอยู่ที่สาขาวิชาที่มีความต้องการสูงในองค์กรภาคเอกชน ซึ่งลักษณะดังกล่าวสอดคล้องกับ

งานวิจัยในอดีตที่ผ่านมา ซึ่งมักพบว่าข้อมูลในตลาดแรงงานที่ได้จากการเก็บข้อมูลทางอินเทอร์เน็ต มีแนวโน้มที่จะไม่สามารถใช้เป็นตัวแทนของประชากร เนื่องจากปัญหา selection bias เช่น ผู้ใช้งานอินเทอร์เน็ตมักจะเป็นกลุ่มที่มีอายุน้อย มีการศึกษาที่สูงกว่าประชากรโดยทั่วไป หรือเป็นผู้ที่ต้องการงานในภาคเอกชนซึ่งมีค่าจ้างที่สูง (Kuhn & Shen, 2013; Kurekova & Zilincikova, 2018)

เพื่อแสดงให้เห็นอย่างชัดเจนถึงข้อควรระวังของการใช้ข้อมูล ORE ในส่วนนี้จึงใช้ข้อมูลผู้ประกันตนรายบุคคล ซึ่งเป็นข้อมูลเพื่อการบริหารจากสำนักงานประกันสังคมที่ครอบคลุม กลุ่มกลุ่มประชากรที่เป็นแรงงานในภาคเอกชน ที่เป็นการจ้างงานที่เป็นทางการ (formal employment) เช่นเดียวกับประชากรของข้อมูล ORE เพื่อมาเปรียบเทียบ รูปที่ 9 แสดงการเปรียบเทียบการกระจายตัวของข้อมูล ORE ตามช่วงอายุ และข้อมูลผู้ประกันตนมาตรา 33 จากสำนักงานประกันสังคม ซึ่งการกระจายตัวของข้อมูลทั้งสองชุดแสดงให้เห็นว่าข้อมูล ORE ที่งานวิจัยนี้ได้สร้างขึ้นมานั้นมีความแตกต่างอย่างชัดเจนจากข้อมูลผู้ประกันตน ซึ่งข้อสรุปดังกล่าวนี้ไม่น่าเป็นที่แปลกใจแต่อย่างใด เมื่อพิจารณาถึงการอภิปรายในหลายส่วนที่ผ่านมา ทั้งยังสอดคล้องกับข้อวิจารณ์ต่อชุดข้อมูลที่เก็บรวบรวมจากทางอินเทอร์เน็ตในงานวิจัยในอดีต

ถึงแม้ว่าข้อมูล ORE ที่ใช้ในงานวิจัยนี้จะมีขนาดใหญ่ แต่นั่นก็ไม่ได้หมายความว่าผู้ใช้ข้อมูลจะสามารถใช้ข้อมูลสถิติทั่วไปที่คำนวณอ้างอิงกับข้อมูล ORE นี้ เพื่อเป็นตัวแทนของประชากรได้ เนื่องจากข้อมูลส่วนใหญ่เป็นข้อมูลที่มาจากกลุ่มที่มีอายุน้อย (20-29 ปี) และมีการกระจุกตัวของกลุ่มตัวอย่างของผู้จบการศึกษาสูงสุดในระดับปริญญาตรี (ตารางที่ 1) และในบางกลุ่มสาขาวิชา (ตารางที่ 8) ซึ่งเป็นปัญหา selection bias ที่ทำให้ข้อมูล ORE ทำให้นักวิจัยจะต้องระมัดระวังหากต้องการประยุกต์ใช้ข้อมูล ORE ในงานวิจัย

5. แนวทางการใช้ประโยชน์ข้อมูล ORE

ถึงแม้ว่าชุดข้อมูล ORE ของงานวิจัยนี้¹⁷ จะยังมีข้อจำกัดอยู่หลายประการ ซึ่งหนึ่งในนั้นก็คือการขาดหายไปของข้อมูลเงินเดือนที่แท้จริง (เนื่องจากหนึ่งใน IJP จะเปิดเผยข้อมูลดังกล่าวเฉพาะกับบริษัทหรือองค์กรที่จ่ายค่าโฆษณาเท่านั้น) ซึ่งทำให้งานวิจัยนี้ยังไม่สามารถแสดงตัวอย่างในการนำไปใช้จริงได้อย่างเต็มที่ แต่อย่างไรก็ตาม คุณลักษณะที่สำคัญของข้อมูลชุดนี้ และยังไม่ได้ทำการอภิปรายเจาะจงลงไปหัวข้อที่ผ่านมา ก็คือ (1) คุณลักษณะตามช่วงเวลา และ (2) คุณลักษณะเชิงพื้นที่ โดยคุณลักษณะตามช่วงเวลานั้น มาจากการที่แบบฟอร์มข้อมูลประวัติออนไลน์ในแต่ละ IJP ได้สร้างขึ้นเพื่ออำนวยความสะดวกในการกรอกข้อมูลให้กับผู้ที่เข้ามาใช้บริการฝากประวัติออนไลน์นั้น จะมีข้อมูลปีที่ผู้ใช้บริการจบการศึกษา และช่วงเวลา (เดือน-ปี) ที่ผู้ใช้บริการเคยทำงานหรือกำลังทำงานอยู่ในแต่ละบริษัท ซึ่งจะทำให้เราได้ข้อมูลในรูปแบบ individual panel data ซึ่งเป็นโครงสร้างข้อมูลที่เป็นประโยชน์และเป็นที่ต้องการเป็นอย่างมากสำหรับการวิเคราะห์ โดยเฉพาะอย่างยิ่งสำหรับนักเศรษฐศาสตร์ เนื่องจากข้อมูลแบบ panel data นั้น สามารถช่วยให้ผู้วิจัยสามารถควบคุม individual-specific unobservable effects¹⁸ (Hausman & Taylor, 1981) ได้ ซึ่งทำลดปัญหาจาก sample selectivity หรือ omitted variables bias ได้

¹⁷ จำนวนข้อมูลที่ใช้ในงานวิจัยนี้มาจาก IJP สองรายในตลาดจัดหางานออนไลน์ของประเทศไทย ซึ่งคิดเป็นเพียงหนึ่งในห้าของจำนวนข้อมูลดิบ ซึ่งผู้วิจัยได้ทำการดึงข้อมูลจาก IJP รายอื่นในประเทศไทยและอยู่ในขั้นตอนการพัฒนา

¹⁸ Individual-specific unobservable effects หมายถึง ปัจจัยหรือตัวแปรใดๆ ในระดับปัจเจกบุคคล ที่อาจส่งผลต่อผลลัพธ์หรือปัจจัยที่กำลังศึกษา แต่ไม่สามารถสังเกตหรือเก็บข้อมูลได้

ตารางที่ 9: จำนวนตัวอย่างในชุดข้อมูล balanced panel data ที่สามารถสร้างได้จากข้อมูล ORE แยกตาม Birth Cohort และระยะเวลาหลังจากที่จบการศึกษาสูงสุด (หน่วย: คน)

Birth Cohort	ระยะเวลาหลังจากจบการศึกษาสูงสุด							
	5 ปี	6 ปี	7 ปี	8 ปี	9 ปี	10 ปี	15 ปี	20 ปี
2500	29	31	32	31	30	31	31	34
2501	33	33	34	33	33	34	30	30
2502	59	55	56	54	55	53	52	49
2503	72	69	70	68	66	67	71	64
2504	62	66	72	75	72	73	75	72
2505	118	121	121	121	126	122	119	102
2506	128	123	129	133	136	145	144	121
2507	190	188	188	190	190	185	169	158
2508	251	244	251	258	263	256	234	218
2509	329	332	331	330	333	345	323	295
2510	443	439	449	452	445	452	408	332
2511	587	585	590	590	590	583	534	413
2512	664	675	684	686	675	673	601	413
2513	819	837	825	848	843	851	756	430
2514	1,015	1,038	1,066	1,085	1,086	1,098	914	501
2515	1,273	1,290	1,311	1,361	1,359	1,338	1,128	505
2516	1,491	1,509	1,535	1,539	1,538	1,554	1,185	496
2517	2,003	2,042	2,098	2,106	2,085	2,054	1,390	491
2518	2,572	2,616	2,608	2,597	2,570	2,521	1,478	450
2519	3,371	3,436	3,436	3,423	3,322	3,177	1,567	395
2520	4,515	4,596	4,571	4,498	4,267	4,004	1,743	354
2521	5,862	5,817	5,681	5,447	5,016	4,512	1,770	309
2522	7,434	7,374	7,093	6,558	5,842	5,006	1,762	214
2523	9,388	9,147	8,479	7,514	6,306	5,254	1,514	153
2524	10,728	10,016	8,838	7,488	6,186	5,069	1,203	91
2525	11,910	10,383	8,640	7,085	5,754	4,603	871	54
2526	12,247	10,298	8,360	6,776	5,348	4,215	606	28
2527	11,906	9,688	7,790	6,166	4,841	3,663	398	23

ตารางที่ 9: จำนวนตัวอย่างในชุดข้อมูล balanced panel data ที่สามารถสร้างได้จากข้อมูล ORE แยกตาม Birth Cohort และระยะเวลาหลังจากที่จบการศึกษาสูงสุด (ต่อ) (หน่วย: คน)

Birth Cohort	ระยะเวลาหลังจากจบการศึกษาสูงสุด							
	5 ปี	6 ปี	7 ปี	8 ปี	9 ปี	10 ปี	15 ปี	20 ปี
2528	11,179	8,887	7,029	5,471	4,114	3,027	310	11
2529	9,897	7,847	6,041	4,589	3,307	2,199	166	5
2530	8,811	6,866	5,237	3,718	2,535	1,642	127	NA
2531	7,536	5,739	4,128	2,779	1,788	1,150	49	NA
2532	7,017	5,042	3,369	2,070	1,325	829	31	NA
2533	5,999	4,149	2,549	1,609	982	586	17	NA
2534	4,352	2,920	1,891	1,209	667	347	9	NA
2535	3,060	2,120	1,360	792	392	232	2	NA
2536	2,026	1,441	828	422	214	119	2	NA
2537	1,443	916	490	258	142	65	NA	NA
2538	819	488	258	134	56	24	NA	NA
2539	394	249	109	53	28	11	NA	NA
2540	198	132	67	34	12	8	NA	NA
จำนวนตัวอย่างรวม	152,230	129,844	108,694	90,650	74,939	62,177	21,789	6,811

ตารางที่ 9 แสดงจำนวนตัวอย่างแบบ balanced panel data โดยทำการดึงข้อมูลจากตัวอย่างที่มีช่วงเวลาของข้อมูลตรงกัน เช่น ตัวอย่าง A ระบุข้อมูลการทำงานอยู่ในช่วงมิถุนายน 2546 จนถึงธันวาคม 2549 ในขณะที่ ตัวอย่าง B ระบุข้อมูลการทำงานอยู่ในช่วงมิถุนายน 2547 จนถึงธันวาคม 2550 ดังนั้น หากเราต้องการจะสร้างชุดข้อมูลแบบ balanced panel data ข้อมูลที่เราสามารถใช้ได้ก็คือ ช่วงเดือนมิถุนายน 2547 จนถึงธันวาคม 2549 เท่านั้น ดังนั้น หากเราต้องการหาจำนวนตัวอย่างในลักษณะ balanced panel data ของกลุ่มการเกิด (birth cohort) ในปี พ.ศ. 2500 และสมมติว่าผู้วิจัยต้องการศึกษาผลตอบแทนทางการศึกษาหลังจากที่ตัวอย่างจบการศึกษามาแล้วห้าปี ซึ่งตัวอย่างที่มีลักษณะดังกล่าวก็มีจำนวนเท่ากับ 29 ตัวอย่าง (ช่องสีเทาใน **ตารางที่ 9**) ซึ่งหากผู้วิจัยต้องการใช้ข้อมูลที่คล้ายกันคือตัวอย่างจบการศึกษามาแล้วห้าปี แต่มีกลุ่มการเกิดตั้งแต่ พ.ศ. 2500 จนถึง พ.ศ. 2540 ก็จะทำให้เรามีจำนวนตัวอย่างทั้งสิ้น 152,230 และจะมีจำนวนน้อยลงเรื่อยๆ หากผู้วิจัยต้องการเพิ่มช่วงเวลาหลังจากจบการศึกษา Appendix C แสดงจำนวนตัวอย่างในชุดข้อมูล balanced panel data ที่สามารถสร้างได้จากข้อมูล ORE แต่แยกตามปีที่จบการศึกษาในระดับการศึกษาสูงสุด (Grad Cohort) และระยะเวลาหลังจากที่จบการศึกษาสูงสุด เพื่อเป็นอีกหนึ่งตัวอย่างในการประยุกต์ใช้ข้อมูล ORE



รูปที่ 10: ที่ตั้งของสถานศึกษาทั้งหมดที่พบในข้อมูล ORE

หมายเหตุ:

1. จุดสีแดงแสดงถึงที่ตั้งของสถานศึกษาแต่ละแห่ง

ในส่วนของคุณลักษณะเชิงพื้นที่ของข้อมูล ORE นั้น ได้มาจากการที่แบบฟอร์มข้อมูลประวัติออนไลน์ของแต่ละ IIP ได้สร้างขึ้นเพื่ออำนวยความสะดวกในการกรอกข้อมูลให้กับผู้ที่เข้ามาใช้บริการฝากประวัติออนไลน์นั้น จะมีช่องให้ผู้เข้ามาใช้บริการสามารถระบุชื่อสถาบันการศึกษาจากสองระดับการศึกษาสูงสุดของตนเองในส่วนของประวัติการศึกษา และช่องสำหรับระบุชื่อและที่อยู่ของบริษัทหรือองค์กรที่ผู้ใช้บริการเคยทำงานในส่วนของประวัติการทำงาน **รูปที่ 10** แสดงการกระจายตัวของตัวอย่างจากข้อมูล ORE โดยใช้ข้อมูลละติจูด/ลองจิจูดของตำแหน่งที่ตั้งสถานศึกษาของตัวอย่างแต่ละราย รวมมากกว่า 20,000 สถานศึกษาทั่วประเทศ ซึ่งข้อมูลเชิงพื้นที่เหล่านี้ นอกจากจะช่วยให้นักวิจัยสามารถควบคุมปัจจัยที่เกี่ยวข้องกับสถานศึกษาแต่ผู้วิจัยอาจจะไม่สามารถสังเกตได้หรือมีข้อมูลที่ไม่เพียงพอ เช่น คุณภาพของสถาบันการศึกษาแล้ว ข้อมูลเชิงพื้นที่เหล่านี้ก็ยังสามารถช่วยให้นักวิจัยสามารถที่จะทำการศึกษาในประเด็นที่เกี่ยวข้องกับการเคลื่อนย้ายแรงงานได้อีกด้วย (Appendix D แสดงผลค่าประมาณที่อาจเบี่ยงเบนหากไม่มีควบคุมปัจจัยที่เกี่ยวข้องกับสถานศึกษา)

6. ข้อสรุปและแนวทางการพัฒนาเพิ่มเติม

งานวิจัยนี้ทำการใช้ข้อมูลประวัติการศึกษาและประวัติการทำงานในอดีตในระดับบุคคล ซึ่งข้อมูลทั้งหมดเป็นข้อมูลตัวอักษร โดยทำการดึงข้อมูลจากฐานข้อมูล ORE ซึ่งถูกตั้งค่าให้สามารถเข้าถึงได้แบบสาธารณะ จากเว็บไซต์ของ IIP จำนวนสองรายในประเทศไทย โดยหลังจากการผนวกรวมชุดข้อมูลจากเว็บไซต์ผู้ให้บริการทั้งสอง และคัดกรองข้อมูลที่มีความเป็นไปได้ที่จะมาจากผู้ที่กำลังหางานคนเดียวกันออก ทำให้เหลือจำนวนข้อมูลรวมสำหรับการวิเคราะห์ ทั้งหมด 1,324,225 คน โดยในจำนวนนี้ แบ่งออกเป็นชุดข้อมูลที่มีเฉพาะผู้ที่มีประสบการณ์ทำงานที่สามารถนำไปจัดโครงสร้างข้อมูลให้เป็นแบบ unbalanced panel data แบบรายเดือน ได้เป็นจำนวนเท่ากับ 710,788 คน และมีจำนวนข้อมูลทั้งสิ้น 35,924,325 คน-เดือน (man-month)

โดยข้อมูล ORE ชี้ให้เห็นว่ารูปแบบการเลือกสาขาวิชาในการศึกษาระดับปริญญาตรี ของแรงงานไทยนั้น ไม่มีการเปลี่ยนแปลงนักในช่วง 40 ปีที่ผ่านมา โดยส่วนใหญ่ก็ยังนิยมที่จะเลือกทำการศึกษาคณะในกลุ่มสังคมศาสตร์และบริหารธุรกิจและคณะวิศวกรรมศาสตร์ เช่นเดียวกันกับอาชีพที่เป็นที่นิยม ก็ยังคงเป็นอาชีพวิศวกร การขายและการตลาด การบุคคล การบัญชี การเงินและธนาคาร เศรษฐศาสตร์ และกลุ่มอาชีพธุรการและการจัดการทั่วไป โดยที่กลุ่มวิชาและกลุ่มสาขาอาชีพที่เกี่ยวข้องกับเทคโนโลยีก็เริ่มได้รับความสนใจมากขึ้นอย่างต่อเนื่อง นอกจากนี้ ถึงแม้ว่าคนรุ่นใหม่จะมีแนวโน้มที่จะไม่ทำงานในองค์กรใดองค์กรหนึ่งนานๆ คนในรุ่นก่อนหน้าก็มีแนวโน้มที่จะปรับตัวตามสถานการณ์ที่เปลี่ยนแปลงไปในตลาดแรงงาน

คุณลักษณะต่างๆ ของข้อมูล ORE ชี้ให้เห็นถึงข้อควรระวังในการประยุกต์ใช้ข้อมูล เนื่องจากข้อมูลส่วนใหญ่เป็นข้อมูลที่มาจากกลุ่มที่มีอายุน้อย (20-29 ปี) และมีการศึกษาที่สูงกว่าประชากรโดยทั่วไป อีกทั้งยังอาจจะมีปัญหาการแบ่งประเภทอาชีพ เนื่องจากชื่อตำแหน่งที่ใช้ในแต่ละองค์กรธุรกิจมีความแตกต่างกันออกไป ซึ่งทำให้ผู้วิจัยจะต้องพัฒนาอัลกอริทึมเพิ่มขึ้นมาใหม่ เพื่อเป็นเครื่องมือที่จะนำมาวิเคราะห์ความเป็นไปได้ในการจัดประเภท หรือชนิดของงานที่ระบุในข้อมูล ORE อย่างไรก็ตาม ด้วยจำนวนของกลุ่มตัวอย่างที่มีขนาดใหญ่ จึงเป็นโอกาสให้นักวิจัยสามารถใช้ประโยชน์จากข้อมูลที่มีความครอบคลุมในเชิงพื้นที่และช่วงเวลา รวมถึงรายละเอียดข้อมูลการศึกษาและประวัติการทำงานที่มีอยู่ในชุดข้อมูล ORE เพื่อเพิ่มประสิทธิภาพในการออกแบบการทดลองกับข้อมูลเพื่อเป็นการลด selection bias ที่เกิดจากการกระจุกตัวของบางสาขาวิชาหรือบางสาขาอาชีพ โดยเฉพาะอย่างยิ่งกับการนำไปใช้กับการศึกษาที่เป็นเชิงสาเหตุ (causal inference studies)

โดยปัญหาที่สำคัญและยังคงอยู่ในขั้นตอนการพัฒนาก็คือ (1) การจัดการกับข้อมูลรายชื่อและที่อยู่ของบริษัทซึ่งกลุ่มตัวอย่างเคยหรือกำลังทำงานอยู่ ซึ่งพบปัญหาเช่นเดียวกันกับชื่อของสถานศึกษาก็คือ การสะกดที่ผิดพลาด แต่เนื่องจากจำนวนของบริษัทหรือองค์กรภาคเอกชนมีจำนวนมากกว่าจำนวนของสถานศึกษาเป็นอย่างมาก ทำให้การจัดการข้อมูลส่วนดังกล่าวเป็นไปอย่างยากลำบาก โดยแนวทางของการแก้ปัญหาดังกล่าว ก็จะเป็นไปเช่นเดียวกันกับการแก้ปัญหาชื่อของสถานศึกษา ซึ่งใช้ความสามารถของ Places API และ Geocoding API ซึ่งพัฒนาโดย Google ซึ่งสามารถทำความเข้าใจเจตนาของคำที่ผู้ค้นหาต้องการจะค้นหาได้ ถึงแม้ว่าจะมีการสะกดคำที่ผิดก็ตาม (2) การจัดการกับข้อมูลชื่ออาชีพและการจัดกลุ่มชื่ออาชีพเพื่อเข้าสู่กลุ่มสาขาอาชีพหลักเหล่านี้ เนื่องจากในความเป็นจริงนั้น ถึงแม้ว่าจะมีหน้าที่หรือมีลักษณะการทำงานที่คล้ายกัน แต่ชื่อเรียกตำแหน่งในแต่ละบริษัทก็จะแตกต่างกันไป เช่น “sales executives” “sales” “ที่ปรึกษาการขาย” “พนักงานขาย” “account executives” ซึ่งชื่อเหล่านี้ทำหน้าที่ที่คล้ายกัน คือ พนักงานขาย แต่ด้วยชื่อที่แตกต่างกัน ทำให้ผู้วิจัยจะต้องทำ corpus ของชื่อตำแหน่งซึ่งบางตำแหน่งมีการใช้ชื่อในรูปแบบที่แตกต่างกันเป็นจำนวนมาก นอกจากนี้ ยังพบปัญหาเรื่องการสะกดคำผิด ซึ่งทำให้ต้องมีการพัฒนาอัลกอริทึมขึ้นมาเพื่อแก้ปัญหานี้โดยเฉพาะ

นอกจากนี้ การพัฒนาข้อมูล ORE สำหรับงานวิจัยในอนาคต ยังรวมถึงการผนวกรวมชุดข้อมูล ORE ขนาดใหญ่ซึ่งผู้วิจัยได้ทำการดึงข้อมูลจากเว็บไซต์ผู้ให้บริการจัดหางานขนาดใหญ่ของประเทศไทยจำนวนสามราย ซึ่งไม่มีการปกปิดข้อมูลค่าจ้างเงินเดือนและยังมีขนาดของฐานข้อมูลใหญ่ ซึ่งทำให้ในขณะนี้ ผู้วิจัยมีคลังข้อมูลดิบเพิ่มเติมอีกกว่า 7 ล้าน user id เพื่อนำมาปรับปรุงให้จำนวนตัวอย่างของข้อมูล ORE นี้สามารถครอบคลุมแรงงานในประเทศไทยให้ได้มากที่สุด

7. บรรณานุกรม

- Acemoglu, D. and Pischke, J.S., 2001. Changes in the wage structure, family income, and children's education. *European Economic Review*, 45(4-6), pp.890-904.
- Ahn, H.J. and Hamilton, J.D., 2020. Measuring labor-force participation and the incidence and duration of unemployment (No. w27394). National Bureau of Economic Research.
- Apaitan, T. and Wibulpolprasert, W., 2018. Stylized Facts on Thailand's Residential Electricity Consumption: Evidence from the Provincial Electricity Authority (No. 107). Puey Ungphakorn Institute for Economic Research.
- Askitas, N. and Zimmermann, K. F., 2015. The internet as a data source for advancement in social sciences. *International Journal of Manpower*.
- Attavanich, W., S. Chantararat and B. Sa-ngimnet, 2018. Microscopic view of Thailand's agriculture through the lens of farmer registration and census data. Forthcoming PIER Discussion Paper.
- Besley, T., Folke, O., Persson, T. and Rickne, J., 2017. Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden. *American economic review*, 107(8), pp.2204-42.
- Böckerman, P., Haapanen, M. and Jepsen, C., 2018. More skilled, better paid: labour-market returns to postsecondary vocational education. *Oxford Economic Papers*, 70(2), pp.485-508.
- Brown, J. and Matsa, D.A., 2016. Boarding a sinking ship? An investigation of job applications to distressed firms. *The Journal of Finance*, 71(2), pp.507-550.
- Brown, J. and Matsa, D.A., 2019. Locked in by leverage: Job search during the housing crisis. *Journal of Financial Economics*.
- Cajner, T. and Ratner, D., 2016. A cautionary note on the help wanted online data. FEDS Notes, Board of Governors of the Federal Reserve System <https://www.federalreserve.gov/econresdata/notes/feds-notes/2016/acautionary-note-on-the-help-wanted-online-data-20160623>. Html.
- Chetty, R., Hendren, N., Kline, P. and Saez, E., 2014. Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, 129(4), pp.1553-1623.
- Chetty, R., 2012. Time trends in the use of administrative data for empirical research. Presentation slides. http://www.rajchetty.com/chettyfiles/admin_data_trends.pdf
- Deming, D. and Kahn, L. B., 2018. Skill requirements across firms and labor markets: Evidence from job postings for

- professionals. *Journal of Labor Economics*, 36(S1), pp.S337-S369.
- Einav, L. and Levin, J., 2014. Economics in the age of big data. *Science*, 346(6210).
- Freeman, R.B., 2002. The labour market in the new information economy. *Oxford Review of Economic Policy*, 18(3), pp.288-305.
- Hausman, J.A. and Taylor, W.E., 1981. Panel data and unobservable individual effects. *Econometrica: Journal of the Econometric Society*, pp.1377-1398.
- Hershbein, B. and Kahn, L. B., 2018. Do recessions accelerate routine-biased technological change? Evidence from vacancy postings. *American Economic Review*, 108(7), pp.1737-72.
- Hummels, D., Jørgensen, R., Munch, J. and Xiang, C., 2014. The wage effects of offshoring: Evidence from Danish matched worker-firm data. *American Economic Review*, 104(6), pp.1597-1629.
- Juhn, C. and Murphy, K.M., 1997. Wage inequality and family labor supply. *Journal of labor economics*, 15(1, Part 1), pp.72-97.
- Kleven, H.J., Landais, C., Saez, E. and Schultz, E., 2014. Migration and wage effects of taxing top earners: Evidence from the foreigners' tax scheme in Denmark. *The Quarterly Journal of Economics*, 129(1), pp.333-378.
- Koenig, F., Manning, A. and Petrongolo, B., 2016. Reservation wages and the wage flexibility puzzle.
- Koijen, R., Van Nieuwerburgh, S. and Vestman, R., 2014. Judging the quality of survey data by comparison with "truth" as measured by administrative records: evidence from Sweden. In *Improving the Measurement of Consumer Expenditures* (pp. 308-346). University of Chicago Press.
- Krueger, A.B., Mas, A. and Niu, X., 2017. The evolution of rotation group bias: Will the real unemployment rate please stand up?. *Review of Economics and Statistics*, 99(2), pp.258-264.
- Kuhn, P. and Shen, K., 2013. Gender discrimination in job ads: Evidence from china. *The Quarterly Journal of Economics*, 128(1), pp.287-336.
- Kurekova, L. M. & Zilincikova, Z., 2018. What is the value of foreign work experience for young return migrants?. *International Journal of Manpower*.
- Lekfuangfu, N., Nakavachara, V., and Sawaengsuksant, P., 2017. Two sides of the same coin: Glancing at labour market mismatch with user-generated internet data. *PIER Discussion Paper No.53*.
- Mamertino, M. and Sinclair, T.M., 2016. Online Job Search and Migration Intentions Across EU Member States. *Institute for International Economic Policy Working Paper Series*.
- Marinescu, I., 2017. The general equilibrium impacts of unemployment insurance: Evidence from a large online job board. *Journal of Public Economics*, 150, pp.14-29.
- Marinescu, I. and Rathelot, R., 2018. Mismatch unemployment and the geography of job search. *American Economic*

Journal: Macroeconomics, 10(3), pp.42-70.

Meghir, C. and Palme, M., 2005. Educational reform, ability, and family background. *American Economic Review*, 95(1), pp.414-424.

Peracchi, F. and Welch, F., 1995. How representative are matched cross-sections? Evidence from the Current Population Survey. *Journal of Econometrics*, 68(1), pp.153-179.

Shen, K. and Kuhn, P., 2013. Do Chinese employers avoid hiring overqualified workers? Evidence from an internet job board. In *Labor Market Issues in China*. Emerald Group Publishing Limited.

Sullivan, D. and Von Wachter, T., 2009. Job displacement and mortality: An analysis using administrative data. *The Quarterly Journal of Economics*, 124(3), pp.1265-1306.

Turrell, A., Speigner, B.J., Djumalieva, J., Copple, D. and Thurgood, J., 2019. Transforming naturally occurring text data into economic statistics: The case of online job vacancy postings (No. w25837). National Bureau of Economic Research.

Wasi, N., Paweenawat, S., Ayudhya, C.D.N., Treeratpituk, P. and Nittayo, C., 2019. Labor Income Inequality in Thailand: the Roles of Education, Occupation and Employment History (No. 117). Puey Ungphakorn Institute for Economic Research.

Woollard, M., 2014. 3.1 Administrative Data: Problems and Benefits. A perspective from the United Kingdom. A. Duşa, D. Nelle, G. Stock, G. Wagner (Eds.), *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences*, SCIVERO, Berlin.

วีระชาติ กิเลนทอง และ กิตติพงษ์ เรือนทิพย์, 2560. ข้อเท็จจริงและมายาคติเกี่ยวกับการว่างงานและอัตราค่าจ้างของผู้จบปริญญาตรีและอาชีวศึกษาในประเทศไทย. /https://www.pier.or.th/wp-content/uploads/2017/05/aBRIDGEd_2017_008.pdf

APPENDIX

Appendix A: คุณลักษณะของข้อมูล ORE ในกลุ่มตัวอย่างผู้จบการศึกษาใหม่

คุณลักษณะ	รวมทุกระดับ	ป.เอก	ป.โท	ป.ตรี	ปวส.
เพศหญิง (%)	71	70	68	73	58
สถานภาพสมรส: โสด (%)	79	79	68	79	77
สถานภาพสมรส: สมรส (%)	1	6	1	1	3
อายุ: ณ ปี 2562 (ปี)	27.18	35.45	32.09	27.65	25.72
อายุ: ณ วันที่แก้ไขข้อมูลล่าสุด (ปี)	22.93	31.73	26.80	23.33	21.67
อายุ: ณ ปีที่จบการศึกษา (ปี)	22.42	31.36	26.32	22.83	21.13
เงินเดือนที่คาดหวัง (บาท/เดือน)	12,613	31,955	18,571	13,076	10,403
จำนวนตัวอย่าง	244,460	33	2,545	199,113	26,306

คุณลักษณะ	รวมทุกระดับ	ปวช.	ม.6	ม.3	ป.6
เพศหญิง (%)	71	66	73	60	38
สถานภาพสมรส: โสด (%)	79	74	82	72	77
สถานภาพสมรส: สมรส (%)	1	3	2	6	8
อายุ: ณ ปี 2562 (ปี)	27.18	23.86	22.79	22.42	23.38
อายุ: ณ วันที่แก้ไขข้อมูลล่าสุด (ปี)	22.93	19.75	19.37	18.59	19.46
อายุ: ณ ปีที่จบการศึกษา (ปี)	22.42	19.13	18.75	17.90	18.77
เงินเดือนที่คาดหวัง (บาท/เดือน)	12,613	9,698	9,534	9,211	10,577
จำนวนตัวอย่าง	244,460	5,532	10,436	482	13

หมายเหตุ:

- คุณลักษณะเพศหญิงและสถานภาพสมรส แสดงข้อมูลเป็นสัดส่วน คุณลักษณะอายุ ประสบการณ์การทำงาน และ เงินเดือนที่คาดหวัง แสดงข้อมูลเป็นค่าเฉลี่ย
- ไม่พบตัวอย่างที่ไม่มีข้อมูล (missing data)
- เนื่องจากมีความเป็นไปได้ที่ตัวอย่างในชุดข้อมูล ORE อาจจะไม่ได้นับที่ข้อมูลประวัติการทำงาน ถึงแม้ว่าจะเคยหรือมีงานทำในขณะที่กรอกข้อมูลประวัติการทำงานออนไลน์ ด้วยเหตุนี้ นิยามของ “ผู้จบการศึกษาใหม่” ของกลุ่มตัวอย่างที่แสดงในตารางนี้ ถูกกำหนดให้ต้องมีคุณสมบัติดังต่อไปนี้
 - ผลต่างระหว่างปีที่มีการแก้ไขข้อมูลล่าสุดและปีที่จบการศึกษาสูงสุด ต้องน้อยกว่าสองปี และ
 - ไม่มีข้อมูลรายละเอียดประวัติการทำงาน (เนื่องจากพบว่า บางตัวอย่างไม่ได้ระบุจำนวนปีประสบการณ์ทำงาน แต่กลับมีการระบุรายละเอียดประวัติการทำงาน) และ
 - จำนวนปีประสบการณ์ทำงานทั้งหมดต้องเท่ากับศูนย์ (เนื่องจาก พบว่าบางตัวอย่างไม่ได้ระบุรายละเอียดประวัติการทำงาน แต่มีการระบุจำนวนปีประสบการณ์ทำงาน)

Appendix B: Keywords สำหรับการค้นหาและจัดกลุ่มสาขาวิชา

<p>สาขาวิชา ศึกษาศาสตร์ (EDU)</p>	<p>ศึกษาศาสตร์ ศึกษา ครุศาสตร์ จัดการประเมินผลทางการศึกษา วิจัยทางการศึกษา เทคโนโลยีและนวัตกรรมการศึกษา ภาวะผู้นำทางการศึกษา จิตวิทยาการศึกษา หลักสูตรการสอน หลักสูตรศึกษา พัฒนาหลักสูตร วิจัยการพัฒนาหลักสูตร สอนคณิตศาสตร์ สอนวิทยาศาสตร์ สอนสังคม สอนภาษาต่างประเทศ สอนภาษาอังกฤษ สอนภาษาฝรั่งเศส สอนภาษาไทย สังคมศึกษา วิทยาศาสตร์ศึกษา สิ่งแวดล้อมศึกษา คณิตศาสตร์ศึกษา คอมพิวเตอร์ศึกษา เทคโนโลยีการศึกษา ธุรกิจศึกษา สุขศึกษา พลศึกษา ศิลปศึกษา ดนตรีศึกษา นาฏศิลป์ศึกษา ศึกษาปฐมวัย ประถมศึกษา มัธยมศึกษา อุดมศึกษา บริหารการศึกษานานาชาติ ครุศาสตร์เกษตร ครุศาสตร์ไฟฟ้า ครุศาสตร์เครื่องกล ครุศาสตร์อุตสาหกรรม ครุศาสตร์เทคโนโลยี อุตสาหกรรมศึกษา ฝึกหัดครู ครูศึกษา อาชีวศึกษา อาชีวศึกษาเทคนิคศึกษา ศึกษาผู้ใหญ่ ศึกษานอกระบบ ศึกษานอกโรงเรียน ศึกษาทางไกล ศึกษาพิเศษ พัฒนศึกษา สารพัดศึกษาหรือพื้นฐานการศึกษา ศึกษาเปรียบเทียบ ศึกษาต่อเนื่อง เทคโนโลยีสื่อสารการศึกษา จิตวิทยาแนะแนว แนะแนวให้คำปรึกษา แนะแนวปฐมวัย นวัตกรรมทางการศึกษา ครุศา</p>
<p>สาขาวิชา มนุษยศาสตร์ (HUM)</p>	<p>ภาษาไทยศึกษา ภาษาไทยตามแนวภาษาศาสตร์ ภาษาไทยปัจจุบัน ภาษาถิ่นไทย ภาษาถิ่นกับสังคมวัฒนธรรมไทย ภาษาไทยในฐานะภาษาต่างประเทศ ศึกษาวินิจฉัย ภาษาไทยสมัยเก่า ภาษาเขมร ภาษาเขมรในฐานะภาษาต่างประเทศ แปลภาษาเขมร วรรณคดีเขมร ภาษาถิ่นกับสังคมวัฒนธรรมเขมร ภาษาลาว ภาษาลาวในฐานะภาษาต่างประเทศ วรรณคดีลาว ภาษาถิ่นกับสังคมวัฒนธรรมลาว ภาษาอังกฤษ ภาษาอังกฤษในฐานะภาษาต่างประเทศ อังกฤษศึกษา ภาษาฝรั่งเศส ภาษาฝรั่งเศสในฐานะภาษาต่างประเทศ ฝรั่งเศสศึกษา ภาษาเยอรมัน ภาษาเยอรมันในฐานะภาษาต่างประเทศ เยอรมันศึกษา วรรณคดีเยอรมัน ภาษาสเปน ภาษาสเปนในฐานะภาษาต่างประเทศ วรรณคดีสเปน สเปนและดินอเมริกา ภาษาอิตาลี ภาษาอิตาลีในฐานะภาษาต่างประเทศ อิตาลีศึกษา วรรณคดีอิตาลี ภาษารัสเซีย ภาษารัสเซียในฐานะภาษาต่างประเทศ รัสเซียศึกษา วรรณคดีรัสเซีย ภาษาโปรตุเกส ภาษาโปรตุเกสในฐานะภาษาต่างประเทศ โปรตุเกสศึกษา วรรณคดีโปรตุเกส ภาษาญี่ปุ่น ภาษาญี่ปุ่นในฐานะภาษาต่างประเทศ ญี่ปุ่นศึกษา วรรณคดีญี่ปุ่น ภาษาบาลีสันสกฤต วรรณคดีบาลีสันสกฤต พุทธศาสนศึกษา ภาษาเวียดนาม เวียดนามในฐานะภาษาต่างประเทศ เวียดนามศึกษา ภาษาเกาหลี ภาษาเกาหลีในฐานะภาษาต่างประเทศ เกาหลีศึกษา วรรณคดีเกาหลี ภาษาถิ่นกับสังคมวัฒนธรรมเกาหลี ภาษาจีน ภาษาจีนในฐานะภาษาต่างประเทศ จีนศึกษา วรรณคดีจีน ภาษาอาหรับ ภาษาอาหรับในฐานะภาษาต่างประเทศ อาหรับศึกษา วรรณคดีอาหรับ ภาษามลายู ภาษามลายูในฐานะภาษาต่างประเทศ วรรณคดีมลายู วรรณคดีเปรียบเทียบ วรรณคดีตะวันตก วรรณคดีตะวันออก วรรณคดีวิจารณ์ วรรณคดีอังกฤษ วรรณคดีอเมริกัน วรรณคดีฝรั่งเศส วรรณคดีไทย วรรณคดีร่วมสมัย วรรณคดีท้องถิ่น วรรณคดีกับสังคมวัฒนธรรมไทย วรรณคดีภาษา วรรณคดีสมัยเก่า ไทยศึกษา ชาติพันธุ์ศึกษา อาณานิคมศึกษา ชุมชนศึกษา วัฒนธรรมศึกษา ศิลปวัฒนธรรมไทย สยามคดีศึกษา ภาษาศาสตร์ สัตวศาสตร์ สัตววิทยา วากยสัมพันธ์ วรรณคดี งานปฏิบัติศาสตร์ ปรัชญา ปรัชญาวิเคราะห์ ภาษาศาสตร์เชิงประวัติ ภาษาศาสตร์สังคม ภาษาศาสตร์จิตวิทยา ภาษาศาสตร์คลังข้อมูล ภาษาศาสตร์คอมพิวเตอร์ บันทึกอนุรักษภาษา ประวัติศาสตร์ ประวัติศาสตร์ไทย ประวัติศาสตร์เอเชียตะวันออกเฉียงใต้ ประวัติศาสตร์จีน ประวัติศาสตร์ญี่ปุ่น ประวัติศาสตร์ตะวันออกกลาง ประวัติศาสตร์โลกมุสลิม ประวัติศาสตร์ยุโรป โบราณคดี เขียนการสร้างบทการแปลบทการดัดแปลงบท วรรณกรรมวิจารณ์และละคร ศาสนาเทววิทยา พุทธศาสนศึกษา อิสลามศึกษา เทววิทยา ศาสนศึกษา ปรัชญา ปรัชญาตะวันตก จริยศาสตร์ อภิปรัชญา ญาณวิทยา สุนทรียศาสตร์ ปรัชญาตะวันออก พุทธปรัชญา ภูมิภาคศึกษา เอเชียศึกษา เอเชียศึกษา ยุโรปศึกษา แอฟริกาศึกษา ลาตินอเมริกาศึกษา พื้นที่ศึกษา เขมรศึกษา อเมริกาศึกษา มุสลิมศึกษา มนุษยศาสตร์ มนุษยศาสตร์ ภาษาฝรั่งเศส รัสเซียศึกษา</p>
<p>สาขาวิชาศิลปกรรม ศาสตร์ (ART)</p>	<p>ทัศนศิลป์ จิตรกรรม ประติมากรรม ภาพพิมพ์ ศิลปะไทย ศิลปะสื่อผสม ศิลปะการจัดวาง ศิลปะกับสิ่งแวดล้อม ภูมิทัศน์ศิลป์ ดิจิทัลอาร์ต คอมพิวเตอร์อาร์ต ภาพยนตร์ ภาพถ่าย ศิลปะภาพถ่าย ศิลปะการถ่ายภาพ อนุภูมิภาคศิลป์ เรขศิลป์ แฟชั่นสิ่งทอ มันชนศิลป์ หัตถศิลป์ นิทรรศการศิลป์ โสตศิลป์ คีตศิลป์ ดุริยางคศิลป์ คีตศิลป์ไทย ดนตรีไทย ดุริยางคศิลป์ไทย ดนตรีสากล ดุริยางคศิลป์ตะวันตก ดนตรีไทยร่วมสมัย ดนตรีสื่อผสม ดนตรีพื้นบ้าน นาฏศิลป์ ศิลปะการแสดง นาฏศิลป์สากล นาฏศิลป์ตะวันตก นาฏศิลป์ตะวันออก นาฏศิลป์ไทย นาฏศิลป์พื้นบ้าน ศิลปะการแสดงร่วมสมัย ศิลปะการแสดงข้ามสื่อ ศิลปะการแสดงผสมสื่อ ศิลปะการละคร ออกแบบเพื่อการแสดง แสดงกำกับการแสดง เพลงดนตรีเพื่อการแสดง ละครประยุกต์ละครเยาวชนละครในการศึกษา ศิลปกรรม ประยุกต์ศิลป์ ออกแบบอุตสาหกรรมศิลป์ ออกแบบผลิตภัณฑ์ ออกแบบเครื่องเคลือบดินเผา ออกแบบเซรามิกส์ ออกแบบแฟชั่น ออกแบบเครื่องประดับ ออกแบบพาณิชศิลป์ ออกแบบนิเทศศิลป์ ออกแบบเลขศิลป์ ออกแบบสื่อโฆษณาภาพยนตร์ ออกแบบสื่อสาร ออกแบบตกแต่งภายใน ออกแบบตกแต่งภายนอก ออกแบบตกแต่งสถาปัตยกรรมภายใน ออกแบบจัดสวน ประวัติศาสตร์ศิลป์ ออกแบบตัดเย็บเสื้อผ้า ศิลปะประยุกต์ ศิลปะประยุกต์ ศิลปะประยุกต์</p>
<p>สาขาวิชานิติศาสตร์ (LAW)</p>	<p>กฎหมายเอกชน กฎหมายมหาชน กฎหมายระหว่างประเทศ กฎหมายสังคม กฎหมายเศรษฐกิจ กฎหมายเปรียบเทียบ นิติปรัชญา กฎหมายธุรกิจ กฎหมายอาญา กฎหมายอาญากระบวนการยุติธรรม กฎหมายอาญาอาชญาวิทยา นิติศาสตร์ กฎหมาย</p>

Appendix B: Keywords สำหรับการค้นหาและจัดกลุ่มสาขาวิชา (ต่อ)

<p>สาขาวิชา สังคมศาสตร์และ บริหารธุรกิจ (SOC)</p>	<p>รัฐศาสตร์ เมืองการปกครอง ปรัชญาทฤษฎีการเมือง เมืองการปกครองเปรียบเทียบ ปกครองท้องถิ่น เมืองระหว่างประเทศ ความสัมพันธ์ระหว่างประเทศ รัฐประศาสนศาสตร์ บริหารรัฐกิจ บริหารทรัพยากรมนุษย์ บริหารองค์การ คลังสาธารณะ นโยบายสาธารณะ บริหารงานท้องถิ่น บริหารเปรียบเทียบ บริหารรัฐกิจบริหารธุรกิจ เศรษฐศาสตร์ เศรษฐศาสตร์ทฤษฎี เศรษฐศาสตร์ปริมาณวิเคราะห์ เศรษฐมิติ เศรษฐศาสตร์ระหว่างประเทศ เงินระหว่างประเทศ ค้าระหว่างประเทศ รวมกลุ่มทางเศรษฐกิจ เศรษฐศาสตร์การคลัง เศรษฐศาสตร์สาธารณะ เศรษฐศาสตร์การเงิน เศรษฐศาสตร์แรงงาน เศรษฐศาสตร์ทรัพยากรมนุษย์ เศรษฐศาสตร์การศึกษา เศรษฐศาสตร์สาธารณสุข เศรษฐศาสตร์สิ่งแวดล้อมทรัพยากรธรรมชาติ เศรษฐศาสตร์สิ่งแวดล้อม เศรษฐศาสตร์ทรัพยากรธรรมชาติ เศรษฐศาสตร์ที่ดิน เศรษฐศาสตร์เกษตร จิตวิทยา จิตวิทยาการทดลอง จิตวิทยาองค์การ จิตวิทยาอุตสาหกรรม จิตวิทยาการเรียนรู้ จิตวิทยาการวิจัย จิตวิทยา จิตวิทยาชุมชน จิตวิทยาเด็กพิเศษ จิตวิทยาผู้สูงอายุ จิตวิทยาพัฒนาการ จิตวิทยาสังคม นิเทศศาสตร์ สื่อสารมวลชน สื่อสารวารสารศาสตร์ ประชาสัมพันธ์ ภาพยนตร์ ภาพถ่าย วิทยุกระจายเสียง วิทยุโทรทัศน์ วาทวิทยา วิธีวิทยาการวิจัย การสื่อสาร สื่อสารการแสดง โฆษณา นิเทศศาสตร์พัฒนาการ พฤติกรรมศาสตร์ ระบาควิทยาสังคม วิจัยพฤติกรรมศาสตร์ ประชากรศาสตร์ วิธีวิทยาการวิจัยทางประชากรศาสตร์ ประชากรกับการพัฒนา เศรษฐศาสตร์ประชากร ประชากรสุขภาพ วิจัย ประชากรสังคม เข้าถิ่น อนามัยการเจริญพันธุ์ ผู้สูงอายุ สมรส เจ็บป่วยตาย สตรีเด็ก สังคมสงเคราะห์ศาสตร์ สังคมสงเคราะห์ สังคม สงเคราะห์ศาสตร์คลินิก สวัสดิการสังคม สวัสดิการชุมชน บริหารนโยบายสวัสดิการสังคม ทฤษฎีการสังคมสงเคราะห์ สังคม สงเคราะห์ทางการแพทย์ สังคมสงเคราะห์ทางการศึกษา แรงงานการพัฒนานโยบายสังคม มานุษยวิทยาวัฒนธรรม มานุษยวิทยา กายภาพ มานุษยวิทยาประยุกต์ สังคมวิทยา มานุษยวิทยา เปลี่ยนแปลงทางสังคมวัฒนธรรม ทฤษฎีสังคม สังคม วิทยาพัฒนาสังคม พัฒนาชุมชน พัฒนาชนบท เพศสภาพศึกษา เพศสถานะศึกษา สตรีศึกษา วิธีวิทยาการวิจัยทางสังคมวิทยาการ ประกันภัย จัดการการเสี่ยงภัยการประกันภัย เศรษฐศาสตร์ประกันภัย บริหารงานยุติธรรม พัฒนาทรัพยากรมนุษย์ รปส ปกครอง บริหารรัฐกิจ ทรัพยากรบุคคล มานุษยวิทยา รัฐประศาสนศาสตร์ บัญชี บัญชีการเงิน บัญชีบริหาร บัญชีต้นทุน ระบบสารสนเทศทางการบัญชี สอบบัญชี ภาษีอากร บริหารธุรกิจ พาณิชยศาสตร์ จัดการ จัดการทรัพยากรมนุษย์ จัดการเชิงกลยุทธ์ จัดการข้อมูลสารสนเทศ ระบบ สารสนเทศเพื่อการจัดการ จัดการการดำเนินงาน จัดการผู้ประกอบการ จัดการธุรกิจระหว่างประเทศ จัดการห่วงโซ่อุปทาน โลจิสติกส์ ภาวะผู้นำพฤติกรรมองค์กร วิเคราะห์ธุรกิจเชิงปริมาณ จัดการเทคโนโลยีนวัตกรรม ธนาคาร จัดการความรู้ ตลาดเงิน บริหารการ ขนส่งระหว่างประเทศ จัดการขนส่งระหว่างประเทศ วิจัยธุรกิจการวิเคราะห์เชิงปริมาณ เงินตลาดหลักทรัพย์ จัดการการผลิต จัดการ อุตสาหกรรม ธุรกิจอสังหาริมทรัพย์ ธุรกิจการเกษตร บริหารธุรกิจบริหารรัฐกิจ บริหารการจัดการ จัดการทั่วไป บริหาร จัดการ business ธุรกิจระหว่างประเทศ management marketing ธุรกิจค้าปลีก อสังหาริมทรัพย์ ค้าปลีก แฟรนไชส์ เลขานุการ ขายส่ง ขายปลีก ขาย บริหารทรัพยากร ธุรกิจประเทศ บรรณารักษศาสตร์ สารสนเทศศึกษา สารสนเทศศาสตร์ จัดการทรัพยากรห้องสมุดบริการ พฤติกรรมสารสนเทศของมนุษย์ จัดการข้อมูลสารสนเทศความรู้ จัดการเอกสารจดหมายเหตุ จัดการองค์กรสารสนเทศ</p>
<p>สาขาวิชา วิทยาศาสตร์ กายภาพ (SCI)</p>	<p>คณิตศาสตร์ทั่วไป ตรรกศาสตร์เชิงคณิตศาสตร์พื้นฐาน ศาสตร์เชิงการจัดคณิตศาสตร์ โทโพโลยี วิเคราะห์เชิงตัวเลข วิทยาศาสตร์ คอมพิวเตอร์ วิจัยดำเนินการ กำหนดการเชิงคณิตศาสตร์ ทฤษฎีเกมส์ ทฤษฎีระบบ คณิตศาสตร์ประยุกต์ จำลองแบบเชิงคณิตศาสตร์ จำลองสถานการณ์ เคมีอินทรีย์ เคมีผลิตภัณฑ์ธรรมชาติ เคมีสังเคราะห์ เคมีการแพทย์ เคมีนาโน เคมีซูพราโมเลกุลาร์ เคมีวิเคราะห์ เคมีอินทรีย์ เคมีเชิงฟิสิกส์ เคมีโลหะอินทรีย์ เคมีสิ่งแวดล้อม เคมีทฤษฎี เคมีคอมพิวเตอร์ เคมีควอนตัม เคมีนิวเคลียร์ ชีวเคมี เคมี ชีววิทยา ชีววิทยา พฤกษศาสตร์ สัตววิทยา วิทยาเชื้อรา เทคโนโลยีการหมัก ฟิสิกส์ ฟิสิกส์เชิงเคมี ชีวฟิสิกส์ ดาราศาสตร์ ฟิสิกส์ ทัศน ศาสตร์ ธรณีฟิสิกส์ ฟิสิกส์เชิงทฤษฎี ฟิสิกส์นิวเคลียร์ ฟิสิกส์พลาสมา ฟิสิกส์ของอนุภาค ฟิสิกส์ควอนตัม ฟิสิกส์เสียง วัสดุศาสตร์ โลหะวิทยา วิทยาการเซรามิกส์ คอมพิวเตอร์ ฟิสิกส์ วิทยาการพอลิเมอร์ วิทยาการสิ่งทอ วิทยาศาสตร์โลก ธรณีวิทยา ปฐพีวิทยา สมุทรศาสตร์ อุทกศาสตร์ อุตุนิยมวิทยา ดาราศาสตร์ ชีวสารสนเทศศาสตร์ นิเวศวิทยา วิทยาศาสตร์สิ่งแวดล้อม วิทยาศาสตร์เพื่อสุขภาพและความ งาม สหวิทยาการเทคโนโลยีและนวัตกรรม วิทยาศาสตร์ทั่วไป สถิติศาสตร์ สถิติ สถิติประยุกต์ ชีวสถิติ คณิตศาสตร์สถิติ วิเคราะห์ สถิติ คณิตศาสตร์ประกันภัย ภูมิศาสตร์ ภูมิศาสตร์เกษตร ภูมิศาสตร์เมือง ภูมิศาสตร์การเมือง ภูมิศาสตร์กายภาพ ภูมิศาสตร์เศรษฐกิจ ภูมิศาสตร์มนุษย์ สถิติทางภูมิศาสตร์ เทคโนโลยีทางภูมิศาสตร์ วิเคราะห์เชิงปริมาณทางภูมิศาสตร์ ภูมิอากาศวิทยา สารสนเทศ ภูมิศาสตร์</p>

Appendix B: Keywords สำหรับการค้นหาและจัดกลุ่มสาขาวิชา (ต่อ)

<p>สาขาวิชา เกษตรศาสตร์ (AGR)</p>	<p>ปรับปรุงพันธุ์พืชไร่ ระบบการปลูกพืชผลผลิตพืชไร่ สรีรวิทยาการผลิตพืชไร่ พืชอาหารสัตว์ พืชศาสตร์ พืชสวน โรคพืชวิทยา กีฏวิทยา พืชวิทยาของแมลง อนุกรมวิธานของแมลง โรควิทยาของแมลง กีฏวิทยาส่งแวดล้อม ปรับปรุงพันธุ์สัตว์ โภชนศาสตร์สัตว์ สัตวศาสตร์ เทคโนโลยีการผลิตสัตว์ เทคโนโลยีการผลิตพืช เทคโนโลยีชีวภาพเกษตร เกษตรกลวิธาน เทคโนโลยีหลังการเก็บเกี่ยว ส่งเสริมการเกษตร รัฐพีศาสตร์ ความอุดมสมบูรณ์ของดิน จุลินทรีย์ดิน สำรวจดิน อนุรักษดิน ฟิสิกส์ของดิน จัดการดิน เคมีของดิน เกษตร จัดการทรัพยากรธรรมชาติสิ่งแวดล้อม จัดการป่าไม้ จัดการลุ่มน้ำ จัดการสิ่งแวดล้อม ชีววิทยาป่าไม้ ความหลากหลายทางชีวภาพ นิเวศวิทยาป่าไม้ พืชวิทยาสิ่งแวดล้อม วนวัฒนวิทยาป่าไม้ เพาะเลี้ยงสัตว์น้ำ เทคโนโลยีการประมง ชีววิทยาประมง อนุกรมวิธานสัตว์น้ำ พันธุศาสตร์สัตว์น้ำ เทคโนโลยีชีวภาพสัตว์น้ำ วาริชศาสตร์ อุตสาหกรรมประมง ประมง</p>
<p>สาขาวิชาบริการ โรงแรมและการ ท่องเที่ยว (SER)</p>	<p>สาธารณสุขศาสตร์ ระบบสุขภาพการบริหารจัดการ โภชนศาสตร์ โภชนาการ โภชนาการชุมชน โภชนาการวัยเด็ก โภชนาการผู้สูงอายุ สาธารณสุข อายุชีวอนามัย ความปลอดภัย อายุชีวอนามัยและความปลอดภัย สุขศาสตร์อุตสาหกรรม เคหกรรมศาสตร์ อาหาร โภชนาการ อาหาร โภชน ผ้าและเครื่องแต่งกาย ท่องเที่ยว การโรงแรม โรงแรมการท่องเที่ยว ท่องเที่ยว พัฒนาแหล่งท่องเที่ยว การขนส่ง โรงแรม ท่องเที่ยว บริการ ธุรกิจ การตลาด และที่ พัก เคหกรรมศาสตร์ ประยุกต์ อุตสาหกรรม งานอาหาร สุขศาสตร์และการบริการ ด้านสุขภาพ อุตสาหกรรม ท่องเที่ยว ธุรกิจ โรงแรม พัฒนา ท่องเที่ยว ธุรกิจ การบิน ธุรกิจ บิน จัดการ ครัวและศิลปะการประกอบอาหาร</p>
<p>สาขาวิชา วิศวกรรมศาสตร์ (ENG)</p>	<p>วิศวกรรมเคมี เคมีวิศวกรรม เคมีเทคนิค เคมีอุตสาหกรรม วิศวกรรมกระบวนการเคมี วิศวกรรมปิโตรเคมี วิศวกรรมชีวภาพ วิศวกรรมเคมีชีวภาพ วิศวกรรมชีวการแพทย์ วิศวกรรมเนื้อเยื่อ วิศวกรรมประสาท วิศวกรรมโยธา ก่อสร้าง วิศวกรรมโครงสร้าง วิศวกรรมโยธา การบริหารการก่อสร้าง วิศวกรรมโยธา โครงสร้างพื้นฐาน วิศวกรรมสิ่งแวดล้อม วิศวกรรมโยธา ชลประทาน วิศวกรรมทรัพยากรน้ำ วิศวกรรมทรัพยากรธรรมชาติ เหมืองแร่ วิศวกรรมทรัพยากรธรณี วิศวกรรมเหมืองแร่ วิศวกรรมสำรวจ วิศวกรรมสำรวจ วิศวกรรมขนส่ง วิศวกรรมสำรวจ วิศวกรรมสำรวจชนบท วิศวกรรมสำรวจ สารสนเทศภูมิศาสตร์ วิศวกรรมการผลิต วิศวกรรมการออกแบบการผลิต วิศวกรรมอุตสาหกรรม วิศวกรรมอุตสาหกรรมระบบ วิศวกรรมอุตสาหกรรม โลจิสติกส์ วิศวกรรมเครื่องกล วิศวกรรมยานยนต์ วิศวกรรมเครื่องกล การออกแบบ วิศวกรรมเครื่องจักรกลเกษตร วิศวกรรมการบินอวกาศ วิศวกรรมอากาศยาน วิศวกรรมทางทะเล วิศวกรรมเรือ วิศวกรรมต่อเรือ วิศวกรรมเครื่องกลเรือ วิศวกรรมไฟฟ้า วิศวกรรมไฟฟ้าสื่อสาร วิศวกรรมไฟฟ้ากำลัง วิศวกรรมอิเล็กทรอนิกส์ วิศวกรรมคอมพิวเตอร์ วิศวกรรมซอฟต์แวร์ วิศวกรรมโทรคมนาคม สื่อสาร วิศวกรรมโทรคมนาคม วิศวกรรมการสื่อสาร วิศวกรรมมาตรวิทยา วิศวกรรมเครื่องมือวัด วิศวกรรมควบคุมอัตโนมัติ วิศวกรรมเครื่องมือมือ วิศวกรรมควบคุมเครื่องมือวัด วิศวกรรมแมคคาทรอนิกส์ วิศวกรรมเกษตร จักรกลเกษตร วิศวกรรมเกษตร วิศวกรรมหลังการเก็บเกี่ยว แปรสภาพ วิศวกรรมอาหาร วิศวกรรมเกษตรอาหาร วิศวกรรมวัสดุ วิศวกรรมวัสดุกระบวนการผลิต วิศวกรรมขนถ่ายวัสดุ วิศวกรรมโลหการ วิศวกรรมพอลิเมอร์ วิศวกรรมเซรามิก วิศวกรรมสิ่งทอ วิศวกรรมเส้นใย วิศวกรรมชีวสารสนเทศ วิศวกรรมศาสตร์ วิศวกรรมเครื่องกล ไฟฟ้า ไฟฟ้ากำลัง ผลิต ออกแบบการผลิต ปิโตรเคมี วิศวกรรมหลังเก็บเกี่ยว แปรสภาพ โยธา ขนถ่าย แมคคาทรอนิกส์ อิเล็กทรอนิกส์ เทคโนโลยีชีวภาพ ชีวสารสนเทศ เทคโนโลยีชีวภาพ วิศวกรรมสิ่งแวดล้อม สร้างสรรค์ เทคโนโลยี สิ่งแวดล้อม สร้างสรรค์ เทคโนโลยี อาหาร เทคโนโลยีพลังงาน เทคโนโลยีสุขภาพ จัดการพลังงาน พลังงานทดแทน เทคโนโลยีสิ่งแวดล้อม เทคโนโลยีการจัดการมลพิษ จัดการสิ่งแวดล้อม ภัยพิบัติ เทคโนโลยีสะอาด ระบบนิเวศ เทคโนโลยีสารสนเทศ สารสนเทศ สื่อสาร โทรคมนาคม เครือข่ายคอมพิวเตอร์ วิทยาการคอมพิวเตอร์ พัฒนาซอฟต์แวร์ พัฒนาระบบ วิทยาการข้อมูล ปฏิสัมพันธ์ระหว่างมนุษย์คอมพิวเตอร์ ระบบสารสนเทศ ความมั่นคงของระบบคอมพิวเตอร์ เทคโนโลยีวัสดุ โลหการ วัสดุ เทคโนโลยีพลาสติก เทคโนโลยีเซรามิกส์ เหมืองแร่ ธรณีวิทยา เหมืองแร่ วิทยาศาสตร์เทคโนโลยีอาหาร วิทยาการอาหาร เทคโนโลยีอาหาร พัฒนาผลิตภัณฑ์อาหาร เทคโนโลยีบรรจุภัณฑ์ นาโนเทคโนโลยี วิทยาศาสตร์นาโน เทคโนโลยีคอมพิวเตอร์ เทคโนโลยีสารสนเทศ information technology วิทยาการคอมพิวเตอร์ คอมพิวเตอร์ เทคโนโลยีสารสนเทศ ธุรกิจ เทคโนโลยีสารสนเทศ เพื่อการออกแบบ เทคโนโลยีมีดเคมี คอมพิวเตอร์แอนิเมชัน วิทยาศาสตร์อาหาร สารสนเทศคอมพิวเตอร์ เทคโนโลยีการประกอบอาหารและการบริการ สถาปัตยกรรมศาสตร์ สถาปัตยกรรม สถาปัตยกรรมไทย สถาปัตยกรรมภายใน ภูมิสถาปัตยกรรม สถาปัตยกรรมผังเมือง เคหการ ออกแบบ ออกแบบนิเทศศิลป์ ออกแบบอุตสาหกรรม ผังเมือง สถาปัตยกรรม</p>

Appendix C: Keywords สำหรับการค้นหาและจัดกลุ่มสาขาอาชีพ

Keywords	ISCO code	นิยามที่ใช้ในงานวิจัยนี้
กฎหมาย/เร่งรัดติดตามหนี้/เร่งรัดหนี้สิน	261/341	LAW
นำเข้า/ส่งออก/shipping/งานด้านต่างประเทศ/จัดซื้อ/คลังสินค้า/logistics	333/335/315/432/511/832/833/835	LOGIS
โรงแรม/ลำม/ไกด์/มัลติเทสต์/งานท่องเที่ยว/งานด้านอาหารเครื่องดื่ม/ภัตตาคาร/บาร์เทนเดอร์/งานบริการทั่วไป/บริการ/พนักงานเสิร์ฟ/สุขภาพความงาม/เสริมสวย/สปา/ฟิตเนส/การ โรงแรม/ท่องเที่ยว	512/513/514/941	TOUR
แพทย์/พยาบาล/เภสัช/สัตวแพทย์/ทันตแพทย์/ผู้ช่วยแพทย์/เภสัชกร/สาธารณสุข	221/222/223/224/225/226/321/322/323/324/325	MED
เกษตร/ปศุสัตว์/ประมง/เหมืองแร่/การเกษตร/การประมง	611/612/613/621/622/631/632/633/634	AGR
สถาปัตยกรรม/ออกแบบตกแต่ง/ประเมินราคา/สถาปัตยกรรมศาสตร์/ตกแต่งภายใน	216	ARCHI
บัญชี/การเงิน/ธนาคาร/หลักทรัพย์/ประกันภัย/ประกันชีวิต/บัตรเครดิต/เศรษฐศาสตร์/สถิติ/งานวิจัย/หุ้น/การประกันภัย/พัฒนาธุรกิจ/พัฒนาองค์กร/iso/งานวิจัย/การเงินการธนาคาร/การธนาคาร/สถิติศาสตร์	212/241/263X/331/332/421/431	ACC
งานการศึกษา/วิชาการ/อาจารย์/ครูศาสตร์/ศึกษาศาสตร์	231/232/233/234/235	EDU
โรงงานการผลิต/ควบคุมคุณภาพ/ช่างยนต์/ช่างกล/เทคนิคอุตสาหกรรม/งานเขียนแบบ/งานdrawing/ช่างไฟฟ้า/ช่างอิเล็กทรอนิกส์/ช่างโลหะ/ช่างแอร์/ไฟร์แมน/ช่างก่อสร้าง/ช่างไม้/ช่างประปา/เจ้าหน้าที่ความปลอดภัยวิชาชีพ/ช่างกลโรงงาน/ช่างเทคนิคอุตสาหกรรม/ช่างสำรวจ/ช่างกลโรงงาน/ช่างเครื่อง/ช่างอิเล็กทรอนิกส์/ช่างทั่วไป/เทคโนโลยีการผลิต/ช่างอื่น	311/312/313/711/712/713/721/722/723/741/742/811/812/813/814/815/816/817/818/821/932	PROD
ผู้จัดการทั่วไป/ผู้อำนวยการ/งานผู้บริหาร	121/122/131/132/133/134/141/142/143	EXEC
นักข่าว/บรรณาธิการ/นักเขียน/แปลภาษา/งานโฆษณา/สิ่งพิมพ์/creative/ช่างภาพ/ศิลปะ/จิตรกรรม/ออกแบบผลิตภัณฑ์/ดีไซน์เนอร์/ช่างตัดเย็บเสื้อผ้า/ช่างแพทเทิร์น/งานบันเทิง/นักแสดง/พรีตตี้/นักดนตรี/นักร้อง/ช่างจิวเวลรี่ /ช่างทองเงิน/ช่างพิมพ์เทคโนโลยีสิ่งพิมพ์/พิมพ์ออฟเซต/ช่างโฆษณา/ช่างพิมพ์ออฟเซต/อิงค์เจ็ท/คหกรรมศาสตร์/เสื้อผ้า/สิ่งทอ/เทคโนโลยีการพิมพ์/นิเทศศาสตร์/สื่อสารมวลชน/ศิลปกรรมศาสตร์	262/263Y/264/265/343/352/731/732/753	ART
วิทยาศาสตร์/เคมี/ชีวะ/ฟิสิกส์/สิ่งแวดล้อม/ธรณี/วิทยาศาสตร์อาหาร/โภชนาการ/ภูมิศาสตร์/ผังเมือง/แผนที่/ดาวเทียม/นักวิทยาศาสตร์การกีฬา/เทรนเนอร์/เทคโนโลยีชีวภาพ/Biotech/วิทยาศาสตร์สิ่งแวดล้อม/อาหาร โภชนาการ/วิทยาศาสตร์การอาหาร	211/314	SCI
พนักงานขับรถ/ส่งเอกสาร/ส่งผลิตภัณฑ์/ติครรถรักษาความปลอดภัย/รถปิกอัพ/ดูแลอาคารสถานที่/แม่บ้าน/พ่อบ้าน/ที่เลี้ยง/ทำความสะอาด/คนงาน/แรงงาน/กรรมกร/ขับรถ	515/911/912/931/933/961/962/921	UNS
ธุรการ/ประสานงาน/คีย์ข้อมูล/ประชาสัมพันธ์/บริการลูกค้า/callcenter/เลขานุการ/คอลเซ็นเตอร์/รัฐศาสตร์/เลขาคอมพิวเตอร์ธุรกิจ/อักษรศาสตร์	334/411/412/413/422/441	ADMIN
การบริหารบุคคล/บุคคล/ฝึกอบรม/training	242	HR

Appendix C: Keywords สำหรับการค้นหาและจัดกลุ่มสาขาอาชีพ (ต่อ)

Keywords	ISCO code	นิยามที่ใช้ในงานวิจัยนี้
เจ้าหน้าที่ขายการตลาด/พนักงานขายการตลาด/ส่งเสริม/สนับสนุน/ประสานงาน การขายการตลาด/การตลาดออนไลน์digitalonlinemarketing/ผู้จัดการขาย/ผู้ช่วย ผู้จัดการขายการตลาด/การตลาด/การขาย	243/522/523/524	MKT
วิศวกรขาย/saleengineer/วิศวกรโยธา/วิศวกรก่อสร้าง/วิศวกรไฟฟ้า/วิศวกรอิเล็กทรอนิกส์/วิศวกรเครื่องกลโรงงาน/วิศวกรอุตสาหกรรม/วิศวกรสื่อสาร/วิศวกร โทรคมนาคม/วิศวกรสิ่งแวดล้อม/วิศวกรเหมืองแร่/วิศวกรรมเคมี/วิศวกรรม สิ่งแวดล้อม/วิศวกรรมสาขา/วิศวกรรมโยธา/วิศวกรรมอุตสาหกรรม/วิศวกรไฟฟ้า/ วิศวกรไฟฟ้า/วิศวกรอิเล็กทรอนิกส์/วิศวกรอิเล็กทรอนิกส์/ วิศวกรรม สื่อสาร/ วิศวกรรมagri/วิศวกรรมสาขาอื่น/วิศวกรรมเกษตร	214/215	ENG
programmer/systemadmin/it/network/ช่างคอมพิวเตอร์/ช่างเทคนิค คอมพิวเตอร์/support/webdesign/graphicdesign/webprogramer/webmaster/วิศวกร คอมพิวเตอร์/วิศวกรระบบ/iosandroid/uxui/admin/socialnetwork/facebook/line/ โปรแกรมเมอร์/วิทยาการคอมพิวเตอร์ วิศวกรรมคอมพิวเตอร์	251/252/351	IT

Appendix D: จำนวนตัวอย่างในชุดข้อมูล balanced panel data ที่สามารถสร้างได้จากข้อมูล ORE แยกตาม Grad Cohort และระยะเวลาหลังจากที่จบการศึกษาสูงสุด

Grad Cohort	ระยะเวลาหลังจากจบการศึกษาสูงสุด							
	5	6	7	8	9	10	15	20
2520	3	3	3	3	3	4	4	5
2521	9	9	9	8	9	10	15	12
2522	12	15	15	16	18	18	20	23
2523	13	13	13	13	15	19	25	28
2524	22	26	29	32	35	34	38	45
2525	34	34	36	36	38	42	58	61
2526	42	46	49	57	60	67	83	86
2527	57	61	65	66	72	78	90	113
2528	77	86	91	96	98	106	120	146
2529	129	141	152	159	167	182	199	230
2530	162	167	181	195	205	216	239	297
2531	186	199	219	227	239	254	272	334
2532	245	267	290	301	308	320	381	424
2533	283	297	312	324	329	346	423	451
2534	370	399	415	438	454	467	586	432
2535	536	545	560	581	608	661	805	587
2536	625	650	676	718	744	773	946	516
2537	746	793	834	903	945	1,007	1,193	512
2538	1,169	1,241	1,337	1,417	1,503	1,618	1,749	659
2539	1,477	1,583	1,692	1,816	1,913	2,055	1,818	639
2540	1,867	2,013	2,185	2,346	2,504	2,660	1,882	582
2541	2,256	2,519	2,766	2,988	3,176	3,270	1,862	402
2542	3,948	4,305	4,642	5,000	5,094	5,222	2,148	202
2543	4,664	5,089	5,429	5,635	5,703	5,549	1,957	8
2544	6,112	6,641	6,939	7,011	6,705	5,543	1,770	NA
2545	8,576	8,969	9,142	8,838	7,252	5,913	1,640	NA
2546	10,422	10,715	10,359	8,371	6,831	5,468	1,001	NA
2547	13,337	12,913	10,411	8,583	6,808	5,213	443	NA

Appendix D: จำนวนตัวอย่างในชุดข้อมูล balanced panel data ที่สามารถสร้างได้จากข้อมูล ORE แยกตาม Grad Cohort และระยะเวลาหลังจากที่จบการศึกษาสูงสุด (ต่อ)

Grad Cohort	ระยะเวลาหลังจากจบการศึกษาสูงสุด							
	5	6	7	8	9	10	15	20
2548	14,598	11,723	9,599	7,630	5,820	4,632	11	NA
2549	14,262	11,517	9,155	6,915	5,542	4,067	NA	NA
2550	13,813	10,852	8,138	6,487	4,655	3,274	NA	NA
2551	11,795	8,921	7,098	5,080	3,500	2,158	NA	NA
2552	11,228	8,949	6,414	4,472	2,648	908	NA	NA
2553	9,427	6,718	4,634	2,700	905	18	NA	NA
2554	8,200	5,700	3,286	1,141	30	NA	NA	NA
2555	7,230	4,309	1,503	52	NA	NA	NA	NA
2556	4,416	1,493	35	NA	NA	NA	NA	NA
จำนวนตัวอย่างรวม	152,348	129,921	108,713	90,655	74,936	62,172	21,778	6,794