



UNIVERSIDAD NACIONAL DE CÓRDOBA

Facultad de Ciencias Económicas, Facultad de Ciencias Agropecuarias y Facultad de Matemática, Astronomía, Física y Computación.

Maestría en Estadística Aplicada

METODOLOGÍAS DE ANÁLISIS LONGITUDINAL PARA DETERMINAR PERFILES DE TRABAJADORES INFORMALES EN GRANDES CENTROS URBANOS DE ARGENTINA.

Tesis para optar al grado académico de Magíster en Estadística Aplicada

Candidato: Lic. Maximiliano Iglesias¹

Directora de Tesis: Dra. María Inés Stimolo²

2021

¹ UNIVERSITY OF CORDOBA, INSTITUTE OF STATISTICS AND DEMOGRAPHY, 5000 CÓRDOBA, ARGENTINA. Email address: miglesias@unc.edu.ar

² UNIVERSITY OF CORDOBA, FACULTY OF ECONOMICS SCIENCE, 5000 CÓRDOBA, ARGENTINA, AND CENTRO DE INVESTIGACIONES EN CIENCIAS ECONÓMICAS, CIECS UNC-CONICET, 5000 CÓRDOBA, ARGENTINA. Email address: maria.ines.stimolo@unc.edu.ar



METODOLOGÍAS DE ANÁLISIS LONGITUDINAL PARA DETERMINAR PERFILES DE TRABAJADORES INFORMALES EN GRANDES CENTROS URBANOS DE ARGENTINA por Maximiliano Iglesias se distribuye bajo una [Licencia Creative Commons Atribución 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/).

COMISIÓN ASESORA DE TESIS

Directora:

Dra. María Inés Stimolo
Facultad de Ciencias Económicas
Universidad Nacional de Córdoba

Tribunal Evaluador:

Dra. Margarita Díaz
Facultad de Ciencias Económicas
Universidad Nacional de Córdoba

Dra. Silvia Ojeda
Facultad de Matemática, Astronomía, Física y Computación
Universidad Nacional de Córdoba

Dr. Pablo Sebastián Gómez
Centro de Investigaciones y Estudios sobre Cultura y Sociedad (CIECS)
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Universidad Nacional de Córdoba

AGRADECIMIENTOS

Quiero expresar mis más sinceros y profundos agradecimientos:

A mi directora de tesis y mentora académica, la Dra. María Inés Stimolo por su acompañamiento a lo largo de toda mi carrera formativa y profesional, por el enorme tiempo dedicado y la constante motivación en cada instancia. Asimismo, por haberme acompañado con este trabajo en congresos, publicaciones y otros espacios de divulgación e intercambio entre pares.

A Gustavo Casanova, Secretario Administrativo de la Escuela de Graduados, que con su constante compromiso nos acompaña y motiva a todos los que transitamos por los distintos cursos y carreras de posgrados.

A la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Córdoba por haber financiado este proyecto, cuyo aporte económico fue indispensable para dedicarle el tiempo necesario al desarrollo de esta investigación.

A los miembros del Tribunal de Tesis: Dra. Margarita Díaz, Dra. Silvia Ojeda y Dr. Pablo Sebastián Gómez por sus importantes aportes y el tiempo dedicado durante las instancias de presentación de este trabajo.

A las personas que forman parte del Instituto de Estadística y Demografía de la Facultad de Ciencias Económicas por haberme brindado un espacio físico, en donde fueron redactadas la mayor parte de estas páginas.

DEDICATORIA

*A mis padres y hermanos,
por los primeros empujones.*

*A Antonella,
por los empujones finales.*

*A mis amigos y amigas en Córdoba,
que con su cariño hicieron que siempre me sienta en mi ciudad.*

RESUMEN

El aumento sostenido de la población con empleo informal en Argentina a partir de la década de los 90' y las condiciones de vulnerabilidad a las que fueron expuestos exhibe la importancia del estudio de factores que inciden sobre trabajadores informales desde un enfoque temporal.

Sin embargo, una limitación para el correcto análisis del mercado laboral en los países en desarrollo, desde un abordaje de la variabilidad temporal, es la falta de información apropiada de datos de panel (Canavire-Bacarreza, Urrego, & Saavedra, 2017), principalmente debido a que los programas estadísticos oficiales presentan paneles muy cortos o corresponden a estudios transversales independientes. De esta manera, se requiere de metodologías que permitan salvar estas dificultades permitiendo incorporar la dimensión temporal como un factor clave de la dinámica y estructura de la problemática objeto de análisis.

Esta investigación tiene por objetivo identificar perfiles de trabajadores informales en los grandes aglomerados urbanos de Argentina durante el período 1990-2003. Esto se logra mediante la combinación de las técnicas de análisis longitudinal: enfoque de pseudo-panel, que posibilita superar la limitación de disponibilidad de datos mediante la construcción de paneles “sintéticos y el método “*KmL*”, que permite identificar la evolución conjunta de trayectorias homogéneas en las cohortes definidas de trabajadores definidas.

Los resultados obtenidos constituyen nuevos aportes en el estudio de la dinámica temporal del mercado de trabajo argentino desde un enfoque no ortodoxo, dado que refuerzan la evidencia de heterogeneidad en la estructura laboral, materializada en la existencia de segmentación socio-demográfica longitudinal.

Palabras clave: Pseudo-panel, K-Means Longitudinal, Segmentación laboral, Informalidad laboral.

ABSTRACT

The sustained increase in the population with informal employment in Argentina since the 1990s and the conditions of vulnerability to which they were exposed shows the importance of studying the factors that affect informal workers from a temporary perspective.

However, a limitation for the correct analysis of the labor market in developing countries, from an approach of temporal variability, is the lack of appropriate information from panel data (Canavire-Bacarreza, Urrego, & Saavedra, 2017), mainly due to the fact that the official statistical programs present very short panels or correspond to independent cross-sectional studies. In this way, methodologies are required to overcome these difficulties, allowing the time dimension to be incorporated as a key factor in the dynamics and structure of the problem under analysis.

This research aims to identify profiles of informal workers in the large urban areas of Argentina during the period 1990-2003. This is achieved through the combination of longitudinal analysis techniques: pseudo-panel approach, which makes it possible to overcome the limitation of data availability through the construction of "synthetic" panels and the "KmL" method, which allows identifying the joint evolution of trajectories homogeneous in the defined cohorts of workers defined.

The results obtained constitute new contributions in the study of the temporal dynamics of the Argentine labor market from a non-orthodox approach, since they reinforce the evidence of heterogeneity in the labor structure, materialized in the existence of longitudinal socio-demographic segmentation.

Keywords: Pseudo-panel, K-Means Longitudinal, Labour market segmentation, Labor informality.

ÍNDICE DE CONTENIDOS

PARTE I. INTRODUCCIÓN.	18
I.1. INTRODUCCIÓN.	19
I.2. PROBLEMA CONCEPTUAL.	24
I.3. PROBLEMA METODOLOGICO.	25
I.4. OBJETIVOS.	27
I.4.1. OBJETIVO GENERAL.	27
I.4.2. OBJETIVOS ESPECIFICOS.	27
PARTE II. MATERIALES.	28
II.1. FUENTES DE DATOS.	29
II.1.1. LAS ENCUESTAS DE HOGARES EN AMÉRICA LATINA.	29
II.1.2. LAS ENCUESTA DE HOGARES DE ARGENTINA.	30
II.1.3. BASES DE DATOS DISPONIBLES.	31
II.2. DEFINICIONES CONCEPTUALES.	32
II.2.1. CENTROS URBANOS.	32
II.2.2. TRABAJADORES ASALARIADOS.	34
II.2.3. EMPLEO ASALARIADO INFORMAL.	34
II.3. PSEUDO PANEL Y VARIABLES ASOCIADAS (POBLACION Y UN. ANALISIS). .	36
II.3.1. COHORTE POBLACIONAL.	36
II.3.2. BASES DE DATOS DE TRABAJO.	38
II.3.3. VARIABLES.	40
II.3.5. RAMA DE ACTIVIDAD ECONÓMICA.	42
II.3.6. INTENSIDAD OCUPACIONAL.	42
II.3.7. EMPRESA O ESTABLECIMIENTO LABORAL.	43
II.3.8. TEMPORAL.	43
II.3.9. INGRESO MONETARIO.	43
PARTE III. ANALISIS EXPLOTARIO DE LAS VARIABLES.	45
III.1. ANÁLISIS DESCRIPTIVO: DATOS AGRUPADOS.	46
III.2. ANÁLISIS DESCRIPTIVO: VARIABLES EN EL TIEMPO.	55

III.2.1. TRABAJO ASALARIADO.....	55
III.2.2. TASA DE DESOCUPACIÓN.....	57
III.2.3. NIVEL EDUCATIVO ALCANZADO.....	59
III.2.4. RAMAS DE ACTIVIDAD ECONÓMICA (CAES).....	60
PARTE IV. METODOLOGÍA.....	67
IV.1. MODELO DE PANEL.....	68
IV.2. MODELO DE PSEUDO-PANEL.....	71
IV.3. METODO DE CLASIFICACIÓN.....	74
PARTE V. RESULTADOS.....	82
V.1. ESTRUCTURA DEL CONJUNTO DE DATOS.....	83
V.2. SOFTWARE Y PAQUETES.....	85
V.3. MODELO DE PSEUDO PANEL: RESULTADOS.....	86
V.4. CLUSTERING LONGITUDINAL.....	94
V.5. ANÁLISIS DE LOS GRUPOS.....	99
PARTE VI. CONCLUSIONES.....	107
VI.1. CONCLUSIONES.....	108
VI.2. LINEAS DE TRABAJO.....	111
PARTE VII. BIBLIOGRAFÍA.....	113
PARTE VIII. ANEXOS.....	125
ANEXO I: INTRODUCCIÓN.....	126
ANEXO II: A. MATERIALES.....	127
ANEXO III: B. ANALISIS EXPLORATORIO.....	136
ANEXO IV: C. METODOLOGÍA.....	145
ANEXO V: D. RESULTADOS.....	146
D.1. RESIDUOS DEL MODELO.....	149
ANEXO VI: E. REPOSITORIO DE DATOS.....	151

ÍNDICE DE FIGURAS

FIGURA III.1. HISTOGRAMA: EMPLEO INFORMAL SEGÚN GÉNERO.....	47
FIGURA III.2. DIAGRAMA DE CAJA Y BRAZOS. DISTRIBUCIÓN DE EMPLEO INFORMAL EN COHORTES SEGÚN AÑO.	48
FIGURA III.3 (A). DIAGRAMA DE DISPERSIÓN. INFORMALIDAD LABORAL Y TAMAÑO DE ESTABLECIMIENTO.	52
FIGURA III.3 (B). DIAGRAMA DE DISPERSIÓN. INFORMALIDAD LABORAL E INTENSIDAD HORARIA.....	52
FIGURA III.3 (C). DIAGRAMA DE DISPERSIÓN. INFORMALIDAD LABORAL Y ANTIGÜEDAD EN LA OCUPACIÓN.....	52
FIGURA III.4. HISTOGRAMA: INTENSIDAD HORARIA Y EMPLEO EN INDUSTRIA SEGÚN GÉNERO.....	54
FIGURA III.5. EVOLUCIÓN DE LA TASA DE DESOCUPACIÓN E INFORMALIDAD EN LOS GRANDES CENTROS URBANOS DE ARGENTINA ENTRE 1990 Y 2003.....	58
FIGURA III.6. PROPORCIÓN DE EMPLEO ASALARIADO EN ACTIVIDADES DEL SECTOR INDUSTRIAL. GRANDES CENTROS URBANOS DE ARGENTINA ENTRE 1990 Y 2003.....	61
FIGURA III.7. PROPORCIÓN DE EMPLEO ASALARIADO EN ACTIVIDADES DEL SECTOR COMERCIO. GRANDES CENTROS URBANOS DE ARGENTINA ENTRE 1990 Y 2003.....	63
FIGURA III.8. PROPORCIÓN DE EMPLEO ASALARIADO EN ACTIVIDADES DEL SECTOR CONSTRUCCIÓN. GRANDES CENTROS URBANOS. ARGENTINA ENTRE 1990 Y 2003.....	64
FIGURA III.9. PROPORCIÓN DE EMPLEO ASALARIADO EN ACTIVIDADES DE LA CONSTRUCCIÓN SEGÚN GÉNERO. GRANDES CENTROS URBANOS DE ARGENTINA. 1990 A 2003.....	65
FIGURA III.10. PROPORCIÓN DE EMPLEO ASALARIADO SEGÚN TIPO DE ESTABLECIMIENTO EMPLEADOR. GRANDES CENTROS URBANOS DE ARGENTINA ENTRE 1990 Y 2003.....	66
FIGURA V.1. ÍNDICE DE CALINSKI Y HARABASZ SEGÚN RE-LANZAMIENTO DEL ALGORITMO. DISTANCIA EUCLIDEANA. ENTRE 2 Y 6 AGRUPAMIENTOS.....	98

FIGURA V.2. CLASIFICACIÓN DE KML3D. TRAYECTORIAS MEDIAS DE VARIABLES INFORMALIDAD LABORAL E INGRESO ESTANDARIZADO PARA CADA CLUSTER. DISTANCIA EUCLIDEANA. 3 CONGLOMERADOS.....	101
FIGURA V.3. DIMENSIONES DE LOS SEGMENTOS LABORALES.....	102
FIGURA V.4. TRAYECTORIAS DE LAS TASAS MEDIAS DE INFORMALIDAD EN LOS CONGLOMERADOS DESAGREGADO SEGÚN GÉNERO.....	103
FIGURA V.5. TRAYECTORIAS DE LOS INGRESOS LABORALES MEDIOS ESTANDARIZADOS EN LOS CONGLOMERADOS DESAGREGADO SEGÚN GÉNERO.....	105
FIGURA I.1. MAPA: PORCENTAJE DE EMPLEO INFORMAL SEGÚN PAÍS.....	126
FIGURA I.2. MAPA: PORCENTAJE DE EMPLEO INFORMAL SEGÚN PAÍS. AMÉRICA DEL SUR.....	126
FIGURA A.1. EVOLUCIÓN POBLACIONAL DE LOS PRINCIPALES CENTROS URBANOS DE ARGENTINA DE 1869 A 2020.....	135
FIGURA A.2. EVOLUCIÓN POBLACIONAL DE LOS PRINCIPALES CENTROS URBANOS DE ARGENTINA DE 1869 A 2020 (EXCLUYENDO EL AGLOMERADO URBANO PARTIDOS DE BUENOS AIRES).....	135
FIGURA B.1. DIAGRAMA DE CAJA Y BRAZOS. DISTRIBUCIÓN DE EMPLEO INFORMAL SEGÚN AÑO.....	137
FIGURA B.2. HISTOGRAMA: ANTIGÜEDAD LABORAL, EMPLEO EN COMERCIO, EMPLEO EN CONTRUCCÓN E INGRESOS LABORALES SEGÚN GÉNERO.....	141
FIGURA B.3 (A). PROPORCIÓN DE ASALARIADOS CON EDUCACIÓN SECUNDARIA COMPLETA. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1990 A 2003.....	142
FIGURA B.3 (B). PROPORCIÓN DE ASALARIADOS CON EDUCACIÓN SECUNDARIA COMPLETA. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1990 A 1994.....	142
FIGURA B.3 (C). PROPORCIÓN DE ASALARIADOS CON EDUCACIÓN SECUNDARIA COMPLETA. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1995 A 1998.....	142
FIGURA B.3 (D). PROPORCIÓN DE ASALARIADOS CON EDUCACIÓN SECUNDARIA COMPLETA. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1999 A 2003.....	142
FIGURA B.4 (A). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD INDUSTRIAL. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1990 A 2003.....	142

FIGURA B.4 (B). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD INDUSTRIAL. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1990 A 1994.....	142
FIGURA B.4 (C). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD INDUSTRIAL. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1995 A 1998.....	142
FIGURA B.4 (D). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD INDUSTRIAL. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1999 A 2003.....	142
FIGURA B.5 (A). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD COMERCIAL. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1990 A 2003.....	143
FIGURA B.5 (B). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD COMERCIAL. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1990 A 1994.....	143
FIGURA B.5 (C). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD COMERCIAL. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1995 A 1998....	143
FIGURA B.5 (D). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD COMERCIAL. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1999 A 2003.....	143
FIGURA B.6 (A). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD CONSTRUCCIÓN. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1990 A 2003.	143
FIGURA B.6 (B). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD CONSTRUCCIÓN. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1990 A 1994.	143
FIGURA B.6 (C). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD CONSTRUCCIÓN. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1995 A 1998.	143
FIGURA B.6 (D). PROPORCIÓN DE ASALARIADOS CON EMPLEO EN ACTIVIDAD CONSTRUCCIÓN. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1999 A 2003.	143
FIGURA B.7 (A). PROPORCIÓN DE ASALARIADOS SEGÚN TAMAÑO DE ESTABLECIMIENTO EMPLEADOR. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1990 A 2003.	144
FIGURA B.7 (B). PROPORCIÓN DE ASALARIADOS SEGÚN TAMAÑO DE ESTABLECIMIENTO EMPLEADOR.. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1990 A 1994.	144

FIGURA B.7 (C). PROPORCIÓN DE ASALARIADOS SEGÚN TAMAÑO DE ESTABLECIMIENTO EMPLEADOR.GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1995 A 1998.	140
FIGURA B.7 (D). PROPORCIÓN DE ASALARIADOS SEGÚN TAMAÑO DE ESTABLECIMIENTO EMPLEADOR. GRANDES CENTROS URBANOS. ARGENTINA. AÑOS 1999 A 2003.	144
FIGURA D.1. TRAYECTORIA DE LA INFORMALIDAD LABORAL EN LAS COHORTES SEGÚN GÉNERO.	146
FIGURA D.2. HISTOGRAMA DE LOS RESIDUOS DEL MODELO. AJUSTE NORMAL (0,1). SUPUESTOS DEL MODELO. NORMALIDAD.	149
FIGURA D.3. QQ-PLOT DE LOS RESIDUOS DEL MODELO (0,1). SUPUESTOS DEL MODELO. NORMALIDAD.....	149
FIGURA D.4. GRÁFICO DE DISPERSIÓN. ORDEN VS RESIDUOS. SUPUESTOS DEL MODELO. INDEPENDENCIA.....	150
FIGURA D.5. GRÁFICO DE DISPERSIÓN. RESIDUOS VS PREDICHOS. SUPUESTOS DEL MODELO. HOMOCEDASTICIDAD.	150

ÍNDICE DE TABLAS

TABLA III.1. CORRELACIÓN ENTRE LA TASA DE INFORMALIDAD Y LAS CARACTERÍSTICAS DE LOS ASALARIADOS EN EL PERÍODO 1990 -2003. MÉTODO "PEARSON'S PRODUCT-MOMENT CORRELATION". BILATERAL.	51
TABLA III.2. CORRELACIÓN ENTRE LA TASA DE INFORMALIDAD Y LAS CARACTERÍSTICAS DE LOS ASALARIADOS EN EL PERÍODO 1990-2003. MÉTODO "PEARSON'S PRODUCT-MOMENT CORRELATION". BILATERAL.	54
TABLA V.1. PRUEBA DE HAUSMAN.....	87
TABLA V.2. MODELO DE PSEUDO-PANEL DE EFECTOS FIJOS (DENTRO DE GRUPOS). PRINCIPALES PRUEBAS DEL MODELO. COEFICIENTES Y ESTADÍSTICOS ASOCIADOS.....	89
TABLA V.3. PRUEBA DE RAZÓN DE VEROSIMILITUD.....	93
TABLA V.4. PROMEDIO DE ÍNDICE DE CALINSKI Y HARABASZ (1974) SEGÚN FUNCIÓN DISTANCIA Y NÚMERO DE GRUPOS SELECCIONADO.....	97
TABLA V.5. COHORTES CONTENIDAS EN CADA CONGLOMERADO. CANTIDAD TOTAL Y DESAGREGADO SEGÚN GÉNERO.....	100
TABLA A.1. PRINCIPALES PROGRAMAS OFICIALES DE ENCUESTAS DE HOGARES EN PAÍSES DE AMÉRICA LATINA.....	127
TABLA A.2. AGLOMERADOS URBANOS RELEVADOS POR ENCUESTA PERMANENTE DE HOGARES EN MODALIDAD PUNTUAL SEGÚN PROVINCIA Y REGIÓN.....	128
TABLA A.3. PRINCIPALES CARACTERÍSTICAS DE LOS AGLOMERADOS URBANOS EN ESTUDIO.....	129
TABLA A.4 COBERTURA POBLACIONAL DE LOS AGLOMERADOS URBANOS EN ESTUDIO.	129
TABLA A.5 (A). COMPOSICIÓN ADMINISTRATIVA Y DEMOGRÁFICA DEL AGLOMERADO URBANO GRAN MENDOZA.	129
TABLA A.5 (B). COMPOSICIÓN ADMINISTRATIVA Y DEMOGRÁFICA DEL AGLOMERADO URBANO PARTIDOS DE BUENOS AIRES.....	130
TABLA A.5 (C). COMPOSICIÓN ADMINISTRATIVA Y DEMOGRÁFICA DEL AGLOMERADO URBANO GRAN TUCUMÁN.	131

TABLA A.5 (D). COMPOSICIÓN ADMINISTRATIVA Y DEMOGRÁFICA DEL AGLOMERADO URBANO GRAN LA PLATA.....	131
TABLA A.5 (E). COMPOSICIÓN ADMINISTRATIVA Y DEMOGRÁFICA DEL AGLOMERADO URBANO GRAN CÓRDOBA.....	132
TABLA A.5 (F). COMPOSICIÓN ADMINISTRATIVA Y DEMOGRÁFICA DEL AGLOMERADO URBANO CIUDAD DE BUENOS AIRES.	133
TABLA A.5 (G). COMPOSICIÓN ADMINISTRATIVA Y DEMOGRÁFICA DEL AGLOMERADO URBANO CIUDAD DE ROSARIO.....	133
TABLA A.6. EVOLUCIÓN DEMOGRÁFICA DE LOS AGLOMERADOS URBANOS EN ESTUDIO SEGÚN REGISTROS CENSALES Y PROYECCIONES OFICIALES. 1869-2020.....	134
TABLA A.7. DICCIONARIO DE VARIABLES DE COHORTES DE TRABAJADORES ASALARIADOS EN BASE DE DATOS DE TRABAJO.	134
TABLA B.1. MEDIDAS DE FORMA EN CARACTERÍSTICAS DE LOS ASALARIADOS EN EL PERÍODO 1990-2003.....	136
TABLA B.2. EDADES DE COHORTES DE TRABAJADORES ASALARIADOS CON EMPLEO INFORMAL EXTREMADAMENTE ALTO SEGÚN AÑO. GRANDES CENTROS URBANOS DE ARGENTINA. PERÍODO 1990-2003.....	136
TABLA B.3. GÉNERO DE COHORTES DE TRABAJADORES ASALARIADOS CON EMPLEO INFORMAL EXTREMADAMENTE ALTO SEGÚN AÑO. GRANDES CENTROS URBANOS DE ARGENTINA. PERÍODO 1990-2003.....	137
TABLA B.4. CÓDIGO DE COHORTES DE TRABAJADORES ASALARIADOS CON EMPLEO INFORMAL EXTREMADAMENTE ALTO SEGÚN AÑO. GRANDES CENTROS URBANOS DE ARGENTINA. PERÍODO 1990-2003.....	137
TABLA B.5. MEDIDAS DESCRIPTIVAS. INFORMALIDAD LABORAL SEGÚN AÑO...	138
TABLA B.6. PRUEBAS DE NORMALIDAD EN CARACTERÍSTICAS DE LOS ASALARIADOS EN EL PERÍODO 1990-2003. BILATERAL.....	139
TABLA B.7. DISTRIBUCIÓN PORCENTUAL DE PERSONAS OCUPADAS SEGÚN CATEGORÍA OCUPACIONAL EN LOS GRANDES CENTROS URBANOS DE ARGENTINA EN EL PERÍODO 1990-2003.....	139

TABLA B.8. PORCENTAJE DE TRABAJADORES ASALARIADOS OCUPADOS EN EMPLEOS INFORMALES Y PORCENTAJE DE TRABAJADORES DESOCUPADOS EN LOS GRANDES CENTROS URBANOS DE ARGENTINA EN EL PERÍODO 1990-2003...	140
TABLA B.9. MEDIDAS DESCRIPTIVAS. VARIABLES EN ASALARIADOS EN EL PERÍODO 1990-2003.....	144
TABLA C.1. CANTIDAD DE PERSONAS RELEVADAS CON EDADES ENTRE 18 Y 60 AÑOS EN LAS ENCUESTAS TRANSVERSALES DE GRANDES CENTROS URBANOS DE ARGENTINA. MUESTRA SIN EXPANDIR.....	145
TABLA C.1. CANTIDAD DE ASALARIADOS RELEVADOS CON EDADES ENTRE 18 Y 60 AÑOS EN LAS ENCUESTAS TRANSVERSALES DE GRANDES CENTROS URBANOS DE ARGENTINA. MUESTRA SIN EXPANDIR.....	145
TABLA E.1. DESCRIPCIÓN DE ARCHIVOS CONTENIDOS EN EL REPOSITORIO DE DATOS.....	152

LISTA DE ABREVIATURAS

AMBA: Área Metropolitana de Buenos Aires.

BT: Base de Trabajo.

BU: Base Usuaría.

BUA: Base Usuaría Ampliada.

CABA: Ciudad de Buenos Aires.

CAES: Clasificador de Actividades Económicas para Encuestas Sociodemográficas.

CEPAL: Comisión Económica para América Latina y el Caribe.

CIET: Conferencia Internacional de Estadígrafos del Trabajo.

CIU: Clasificador Industrial Internacional Uniforme.

DE: Desviación Estándar.

DPE: Direcciones Provinciales de Estadísticas.

EPH: Encuesta Permanente de Hogares.

EPHC: Encuesta Permanente de Hogares modalidad Continua.

EPHP: Encuesta Permanente de Hogares modalidad Puntual.

GBA: Gran Buenos Aires.

IASI: Instituto Interamericano de Estadística.

ILO: International Labour Office (en español OIT).

INDEC: Instituto Nacional de Estadísticas y Censos.

KML: K-Means Longitudinal.

MCG: Mínimos Cuadrados Generalizados.

OIT: Organización Internacional del Trabajo (en inglés ILO).

ONU: Organización de las Naciones Unidas (en inglés UN).

PD: Población Desocupada.

PEA: Población Económicamente Activa.

PREALC: Programa Regional de Empleo para América Latina y el Caribe.

PyMEs: Pequeñas y Medianas Empresas.

TA: Tasa de Actividad laboral.

TD: Tasa de Desocupación laboral.

TE: Tasa de Empleo.

TI: Tasa de Informalidad laboral.

USAID: Agencia de los Estados Unidos para el Desarrollo Internacional.

PARTE I. INTRODUCCIÓN.

En el primer capítulo se introducen los principales elementos conceptuales vinculados al empleo informal, su desarrollo como tópico entre los estudios del mercado de trabajo y su caracterización en Argentina y la región. (I.1). Posteriormente se presentan algunas de las discusiones conceptuales relacionadas a su abordaje (I.2), en donde se enfatiza la importancia de su medición desde un enfoque temporal. Asimismo, se exhiben los problemas metodológicos (I.3) que dificultan este tipo medición y se describen las técnicas que serán desarrolladas para salvar este impedimento. Por último, en la sección I.4 se exponen los objetivos generales y específicos que se abordaran durante el siguiente trabajo.

I.1. INTRODUCCIÓN.

El empleo informal, que se caracteriza por una relación de trabajo desprovista de protección social y legislativa, es un problema que afecta a todos los países. En la actualidad se estima que más del sesenta por ciento de la fuerza laboral del mundo se encuentra trabajando en la economía informal. Esta incluye a todas las actividades económicas de trabajadores o unidades económicas que, en la legislación o en la práctica, no recaen en el ámbito de mecanismos formales o son contempladas de manera insuficiente (OIT, 2018). Sin embargo, su distribución es muy desigual a lo largo del mundo, dado que, los niveles que alcanza exhiben mayor prevalencia en los países en desarrollo. Los informes de la Organización Internacional del Trabajo (OIT)³ indican que, en el año 2016, mientras que en Europa y Asia Central la proporción de empleos informales representaba el 25,1 por ciento; en África la informalidad se ubicaba en torno al 86 por ciento del empleo total. En sintonía con esto, se estima que para ese mismo año, el 93 por ciento del empleo informal en el mundo se encontraba en los países en desarrollo.

También se observaron cambios a lo largo del tiempo. La magnitud de los niveles presentados desde inicio de siglo, y las condiciones de vulnerabilidad que genera en la población alcanzada, manifestaron la importancia del estudio y monitoreo de los conceptos relacionados a este fenómeno.

Sin embargo, las concepciones y terminologías, que fueron aportando conocimiento para la constitución de un marco o conjunto de reglas y principios propios en el estudio de la informalidad como tópico independiente, tienen al menos medio siglo.

³ Organización Internacional del Trabajo (2018). “Mujeres y hombres en la economía informal: un panorama estadístico” (tercera edición). Ginebra: OIT.

En general, los autores especializados en la problemática, coinciden en colocar al informe de la OIT sobre Kenia⁴ (1972) como momento de origen del concepto de “sector informal”. Este documento, junto a otros estudios elaborados por la Oficina Internacional de Empleo y enfocados en países africanos, encuentra su inspiración en el artículo de Keith Hart (1973). Este autor utiliza la definición de sector informal para describir e interpretar la sección en donde eran desarrolladas actividades económicas, de bajos ingresos y con alto grado de informalidad, por parte de los grupos de migrantes del norte del país en las zonas urbanas del sur de Ghana.

Si bien existen algunos antecedentes en años previos⁵, es entonces a partir de esta década que comenzaron a utilizarse, en forma generalizada, los términos vinculados a “sector informal” y “empleo informal”. El primero de estos enfoques complementarios centra su atención en las características de los establecimientos en donde los individuos desempeñan su actividad; mientras que el segundo, hace referencia al incumplimiento legal de los derechos laborales de los individuos en relación a su puesto de trabajo. No obstante, estas definiciones se expresaban, desde muchos aspectos, de una manera todavía poco precisa.

El traslado a América Latina de estas perspectivas, correspondientes a la OIT, estuvo a cargo del Programa Regional de Empleo para América Latina y el Caribe (PREALC) con sede en Chile. A partir de esos años, este centro de estudios asumió la tarea de medir y caracterizar los sectores informales en la región argumentando su origen en su heterogeneidad estructural, la concentración económica y la permanencia de mercados oligopólicos (Neffa, 2016).

⁴ International Labour Office (1972). “Employment, incomes and equality. A strategy for increasing productive employment in Kenya”.

⁵ “El concepto de trabajo o empleo informal fue anticipado en los años 50 por los trabajos académicos de Lewis (1954), sobre las economías en ‘vías de desarrollo’”. Neffa (2009).

El desarrollo y expansión del análisis de este ámbito de estudio del mercado trabajo en este periodo no es un casual. El momento histórico que atravesaban tanto la economía mundial como la correlación de fuerzas entre clases sociales posibilitaron el avance, y en casos la imposición, del paradigma neoliberal. Esta escuela de pensamiento se materializó mediante un conjunto de reformas y políticas económicas que cambiaron el transitar del mundo, en materia de derechos y dinámica laboral, durante el último cuarto del siglo XX.

La crisis de acumulación del sistema capitalista en los años 70 derivó en un proceso de desconcentración y tercerización de la producción junto a las políticas de flexibilización laboral con el fin de reducir los costos y riesgos de producción y evitar la caída de la tasa de ganancia (Castells Manuel y Alejandro Portes, 1986). El impacto generado en la región de América Latina fue profundo. Mientras que, en las pequeñas y medianas empresas de los países capitalistas industrializados, estos mecanismos de flexibilización de fuerza de trabajo, permitieron aumentar su rentabilidad; facilitando acceso a crédito y trabajo calificado. En los países emergentes, estas empresas lo incorporaron como forma de supervivencia permitiéndoles la posibilidad de evadir la legislación y ser una alternativa al desempleo (Neffa, 2009).

Este cambio de paradigma, tanto en la estructura productiva como en la legislación laboral, dio comienzo en América Latina a un ciclo marcado por una tendencia creciente en el número de trabajadores y unidades económicas que desempeñaban sus actividades en un marco de informalidad.

En el caso concreto de la Argentina, lo que actualmente denominamos como economía informal, no representó niveles significativos hasta finalizar la primera mitad de la década mencionada. En general, durante los periodos previos con vigencia del modelo de sustitución de importaciones, el

mercado de trabajo mantuvo niveles de ocupación cercanos al pleno empleo y una extendida protección social (Novick, 2007).

Es a partir del golpe militar de 1976, mediante la imposición de un modelo económico que modificó la matriz productiva nacional dando prioridad a las actividades primarias frente a las industriales, que esta problemática comenzó a tener mayor relevancia. El efecto de las medidas llevadas a cabo, en materia financiera y de comercio exterior, debilitaron al sector de la industria nacional. Estas políticas económicas, junto a las reformas en torno a la ley de contrato que iban en detrimento de los intereses de los trabajadores y que contribuyeron a los procesos de subcontratación y tercerización de actividades, provocaron un crecimiento de la cantidad de trabajadores en el Sector Informal Urbano (Boyer y Neffa, 2007). Este último, funcionó como alternativa al desempleo, subempleo y a las condiciones laborales precarias de ese nuevo escenario del mercado de trabajo argentino.

En la década posterior, a pesar de existir coyunturas en que se contuvieron las políticas de ajuste neoliberales, el deterioro sobre las condiciones de empleo de los trabajadores mantuvo tendencia. Los distintos gobiernos de la región, en función de la correlación de fuerzas presentada, fueron incorporando políticas de ajuste estructural. En general, las reformas económicas que se llevaron a cabo en los países latinoamericanos desde mediados de los 80 y su impacto sobre los mercados de trabajo provocaron un descenso en la calidad del empleo (Weller, 2003). En consecuencia, durante la década del '90, el empleo informal en las áreas urbanas de América Latina creció aproximadamente de un 50% a un 58% (Gasparini y Tornarolli, 2007). Sin embargo, a pesar de que en la década siguiente la región se distinguió por presentar un crecimiento económico acelerado y sostenido con mejoras en las variables macroeconómicas

fundamentales (Kacef & Lopez Monti, 2010), esto no se trasladó con igual intensidad en la calidad del empleo.

En el caso particular de Argentina, el empleo informal creció del 51,1% en 1993 al 59,4% para el segundo trimestre de 2003 (Beccaria & Groisman, 2007). A partir de ese año, junto con el crecimiento de la actividad económica, comenzó a descender hasta estancarse alrededor del 35% en 2008, año en que comenzó a aumentar nuevamente para posicionarse actualmente en torno al 40% de la población ocupada (INDEC, 2018).

La crisis económica y financiera mundial iniciada en el año 2008 ha puesto de manifiesto la importancia del empleo formal, posicionando al estudio de la informalidad laboral en un lugar de relevancia entre los temas de investigación científica relacionados con el desarrollo económico y social, a lo cual se suma la presente investigación.

La condición de informalidad en los trabajadores se caracteriza por tener entre sus consecuencias directas la denegación de derechos laborales esenciales, como lo son la seguridad jurídica y social, la cobertura sanitaria y sindical, entre otros aspectos; mientras que, entre sus efectos indirectos, se destacan la segmentación de la estructura del mercado laboral (Jiménez, 2010), la discriminación por menor remuneración salarial frente a sus pares formales (Maurizio, 2014) y su consecuente influencia en las condiciones socioeconómicas de los hogares que componen, y en relación a lo anteriormente mencionado, es notoria la incidencia de esta diferencial de ingresos sobre los niveles de pobreza (Beccaria y Groisman, 2007).

Cabe aclarar que, esta problemática no se presenta a lo largo del espacio y del tiempo, en igual dimensión, intensidad y magnitud en los diversos ámbitos de incidencia de aquellas personas que conforman al mercado de trabajo. De esta manera, existen ramas de la actividad económica,

como así también, grupos de personas que son más vulnerables a ser alcanzadas por la informalidad, de modo tal que trabajadores muy jóvenes, migrantes, mujeres, minorías étnicas o personas con niveles de educación relativamente bajos aparecen sobre representados en la economía informal (OIT, 2015).

I.2. PROBLEMA CONCEPTUAL.

Por lo expuesto en el apartado anterior, las condiciones de vulnerabilidad a las que son expuestos los trabajadores informales han puesto de manifiesto la importancia del estudio y monitoreo de su evolución a través de tiempo y revela la importancia que tiene el análisis de los múltiples factores que inciden sobre la población con empleo informal desde un enfoque temporal. De esta manera, se requiere de metodologías que permitan incorporar la dimensión temporal como un factor clave para un análisis lo más completo posible de la dinámica y estructura de la problemática objeto de análisis, como así también, la relación con los indicadores macroeconómicos como representantes de las particularidades que presenta cada coyuntura.

Esta investigación incorpora en el área de estudio tres principales aportes. En primer lugar, la medición de los indicadores correspondientes a las dimensiones de informalidad laboral, como son “empleo informal” y “sector informal” para los centros urbanos más importantes de Argentina en el período 1990-2003. En segundo lugar, el análisis de las condiciones sociodemográficas de los trabajadores informales mediante el uso de métodos estadísticos longitudinales que permitan la construcción y el estudio específico de las cohortes de trabajadores según nacimiento y género en estos aglomerados. Por último, la identificación a partir de la evolución conjunta de las trayectorias de sus características sociodemográficas y económicas de grupos de cohortes de trabajadores informales homogéneas.

Cabe destacar que analizar bajo la metodología propuesta los perfiles de las cohortes de trabajadores informales, permitirá establecer de manera más específica los factores que los afectan y la incidencia sobre ellos, de modo que la relevancia de esto último, radica en la posibilidad de generar políticas públicas para mitigar la problemática del empleo informal con mayor nivel de focalización.

Este análisis, tal como se mencionó, se realizará para los principales aglomerados urbanos de Argentina entre los años 1990-2003. Esto incluye Ciudad Autónoma de Buenos Aires, Partidos del Gran Buenos Aires, Gran Rosario, Gran Córdoba, Gran La Plata, Gran Mendoza y Gran Tucumán.

La elección de estos grandes centros urbanos tiene base en criterios concretos que resultan apropiados para mejorar el grado de especificidad en la investigación del fenómeno en estudio. El período fue seleccionado por permitir la distinción de ciclos económicos marcados y bien diferenciados entre sí que permiten y posibilitan percibir la dinámica de las cohortes de los empleados informales en particular; y la incidencia de las reformas sobre las regulaciones laborales y los efectos de factores macroeconómicos en general, constituyéndose un trabajo inédito en cuanto a la metodología y los conceptos a ser utilizados para el caso particular de estos aglomerados de conjunto.

I.3. PROBLEMA METODOLOGICO.

Una limitación para el correcto análisis del mercado laboral en los países en desarrollo desde un abordaje de la variabilidad temporal, es la falta de información apropiada de datos de panel (Canavire-Bacarreza, Urrego, & Saavedra, 2017).

En algunos de estos países se dispone de información de datos de panel provenientes de programas oficiales pero presentan ciertas limitaciones. En encuestas basadas en paneles rotativos, los hogares o individuos permanecen un período relativamente corto en la muestra lo que imposibilita su seguimiento en el tiempo. A su vez, el abandono no aleatorio de ciertas unidades por “desgaste” puede generar un sesgo considerable en las estimaciones (Perera, 2006). Otros países no tienen encuestas de hogares de panel pero cuentan con una serie de secciones transversales independientes provenientes de programas estadísticos realizados durante períodos relativamente largos.

La metodología de pseudo panel permite superar la limitación de disponibilidad de datos mediante la construcción de paneles “sintéticos”. Esto se logra reemplazando las observaciones individuales del panel original con medias de subgrupos de la población, es decir, subgrupos de individuos de los que se puede identificar su aparición en repetidas encuestas transversales (Meng, Brennan, Purshouse & otros, 2014). Este enfoque nos posibilita el seguimiento de cohortes a lo largo del tiempo en secciones transversales repetidas, generando series de tiempo para las medias de los subgrupos que se pueden usar como si fueran datos de panel, para estimar la evolución temporal de las mismas y la construcción de “variables-trayectorias”.

Este trabajo pretende identificar grupos de trabajadores informales con características diferentes que permanecen en el tiempo. Las técnicas de agrupamiento (clúster) estándar permiten identificar grupos, pero no tienen en cuenta el aspecto temporal. Esto implica que deben utilizarse algoritmos que tengan en cuenta la vecindad temporal de los objetos para extraer conocimiento útil para un análisis lo más completo posible de la dinámica y estructura de la problemática objeto de estudio, como así también, la relación entre sus múltiples factores y determinantes. En

este sentido, las técnicas de clúster-longitudinal combinan similitudes de contenido y adyacencia temporal en una sola representación.

El enfoque KmL (K-means Longitudinal) de Genolini y Felissard (2017) permite, por un lado, identificar la evolución conjunta de trayectorias homogéneas en las cohortes definidas a partir del pseudo-panel, posibilitando la conformación de grupos de trabajadores en el periodo analizado. Por otro lado, permite identificar, a partir de la visualización, los cambios de tendencia potenciales factores temporales (reformas, cambios técnicos) que afectan a cada grupo específico.

I.4. OBJETIVOS.

I.4.1. OBJETIVO GENERAL.

- Identificar perfiles de trabajadores informales durante el período 1990-2003 en los grandes aglomerados urbanos de Argentina considerando su evolución temporal a partir de las técnicas de pseudo-panel y K-means Longitudinal.

I.4.2. OBJETIVOS ESPECIFICOS.

- Analizar las características sociodemográficas y la evolución de los factores que inciden en la condición de informalidad de las cohortes de trabajadores en el período 1990 a 2003 en los grandes aglomerados urbanos de Argentina.
- Identificar los factores que influyen en la informalidad de cada uno de los perfiles definidos teniendo en cuenta su trayectoria en los grandes aglomerados urbanos de Argentina en el período 1990-2003 utilizando un pseudo panel.
- Validar la importancia de la incorporación de la dimensión temporal en la construcción de perfiles de trabajadores informales.

PARTE II. MATERIALES.

En el segundo capítulo se describen los materiales que serán utilizados en el análisis de la problemática objeto de estudio. En la primera sección (II.1) se presentan las fuentes de datos secundarios que están disponibles en Argentina y la región, en donde se reflexiona respecto de la calidad de la información recolectada y su evolución a través de los años. En la segunda sección (II.2) se desarrollan y delimitan los principales conceptos que integran el marco de referencia de la presente investigación. Por último, en la sección II.3, se detallan los procedimientos utilizados para la construcción de los pseudo paneles y la agregación de las variables asociadas que conforman la base de datos de trabajo. Esta última, constituye el insumo necesario para la aplicación de las técnicas de análisis longitudinal.

II.1. FUENTES DE DATOS.

II.1.1. LAS ENCUESTAS DE HOGARES EN AMÉRICA LATINA.

La región de América Latina, considerada en conjunto, se ubica en una posición intermedia en cuanto a desarrollo estadístico. Sin embargo, en lo que refiere a calidad de información en estadísticas oficiales, la situación de los países que la integran es muy heterogénea, conviviendo situaciones de alto contraste (Tacla, 2006). El insuficiente grado de cobertura en el registro público nacional, el subregistro en algunas de las subdivisiones administrativas o los errores de contenido en la etapa de administración de cuestionarios son algunos de los problemas que afectan la calidad de las estadísticas en los países de la región (Bay & Orellana, 2007).

Las limitaciones que enfrentan por falta de recursos o escasez en la infraestructura institucional se contraponen con la creciente necesidad de estadísticas de desarrollo actualizadas, fundamentales para el lineamiento y la supervisión de políticas públicas. En consecuencia, las encuestas de hogares por muestreo representan el principal recurso para la obtención actualizada de los indicadores socioeconómicos y demográficos más importantes en los países en desarrollo. En este sentido, es que estas fuentes de información, se constituyeron con el tiempo en el instrumento más idóneo para lograr contrarrestar las deficiencias presentadas (Martine, 2005).

En términos históricos las encuestas de hogares comenzaron su implementación y desarrollo en forma periódica, en América Latina, a inicios de la década de 1960. Los operativos impulsados por La Agencia de los Estados Unidos para el Desarrollo Internacional (USAID) y el Instituto Interamericano de Estadística (IASI) se basaron en el modelo Atlántida⁶ que contenía los elementos metodológicos para la realización de encuestas tal como se habían establecido en los

⁶ US Bureau of the Census. (1969). Atlántida: un estudio de caso en encuestas de hogares por muestra. Series ISPOL. Washington D.C.

países desarrollados. A finales de la década, gran parte de los países de la región, habían constituido programas permanentes de encuestas con el objetivo central de brindar información sobre el mercado de trabajo que sea comparable a nivel continental (ONU, 1983).

En la actualidad los principales países de América Latina mantienen programas permanentes que realizan periódicamente encuestas de hogares, ya sea en modalidad puntual o continua⁷. Estos proyectos han ampliado su cobertura de análisis incorporando información referida a otros campos como educación, vivienda, migración, salud, cultura, entre otras. Sin embargo, son pocos los países que han logrado cobertura nacional urbano-rural estricta.

II.1.2. LAS ENCUESTA DE HOGARES DE ARGENTINA.

En el caso concreto de Argentina, el programa nacional Encuesta Permanente de Hogares (EPH) es el encargado de efectuar la recolección periódica de información referente las condiciones de vida de la población en los principales aglomerados urbanos del país.

El esquema muestral de este operativo se corresponde con un diseño en dos etapas. En la primera etapa se eligen los radios censales mediante el método estratificado. En la segunda etapa se realiza la selección aleatoria para el listado de las viviendas pertenecientes al radio censal extraído. A su vez, cuenta con un esquema de rotación de paneles⁸ que mejora la precisión en las estimaciones (por tamaño y periodos) y reduce el nivel de no respuestas (INDEC, 2003). Esta modalidad de muestreo por aglomerados se aplica desde 1973⁹ y fue incorporando unidades hasta

⁷ La Tabla A.1 del anexo “A. Materiales” muestra los principales programas de encuestas de hogares disponibles en cada uno de los países de América Latina y el Caribe.

⁸ Para conocer respecto del esquema de rotación 2x2x2 se puede consultar en: Instituto Nacional de Estadísticas y Censos (2003). La nueva Encuesta Permanente de Hogares de Argentina. Págs. 19-20.

⁹ Hasta el año 1982, inclusive, la información disponible en base de datos usuarias se limita al Gran Buenos Aires (GBA) que incluye a la Ciudad Autónoma de Buenos Aires (CABA) y partidos del Gran Buenos Aires.

cubrir actualmente 31 aglomerados urbanos y un área urbano-rural (Alto valle)¹⁰. En conjunto, estos agregados, representan más del 70 por ciento de la población urbana de Argentina (Messere & Hoszowski, 2007).

El operativo “en campo” es realizado en forma conjunta por el Instituto Nacional de Estadísticas y Censos (INDEC) y las Direcciones Provinciales de Estadísticas (DPE). Este operativo sufrió un cambio en su relevamiento en el año 2003, pasando de la modalidad puntual con dos ondas anuales, en mayo y octubre, a una de tipo continua con datos trimestrales. Esto implicó el cambio en el dominio de estimación y la incorporación de 8.000 hogares a la muestra, pasando de 17.000 a 25.000 hogares.

El instrumento de recolección consta de tres cuestionarios divididos en bloques: el primero para relevar características de las viviendas, entendidas como el espacio físico donde viven las personas. El segundo corresponde al hogar que es constituido por la persona o el grupo de personas que viviendo bajo el mismo techo comparten los gastos de alimentación. Por último, existe un formulario para cada una de las personas residentes que conforman el hogar. Este último, actualmente solo se realiza en personas de 10 y más años.

II.1.3. BASES DE DATOS DISPONIBLES.

Las bases de datos disponibles para usuarios secundarios, permiten el acceso a una parte importante de la información recolectada en las distintas variables que conforman los cuestionarios. Durante el periodo de la EPH Puntual las bases están provistas en los formatos Base R2 Usuaría (R2), Base Usuaría (BU), Base de Trabajo (BT), Base Usuaría Ampliada

¹⁰ La Tabla A.2 del anexo “A. Materiales” brinda información de todos los aglomerados urbanos relevados por el programa durante el período de referencia del análisis (1990-2003).

(BUA). Mientras que, la información perteneciente a la EPH Continua se dispone en dos componentes: la Base Hogar (Hogar) y la Base Individuos (Individual).

El ciclo en estudio, que comprende el intervalo de años que va desde 1990 a 2003, coincide con la última etapa en que se ejecutó el programa EPH bajo la modalidad puntual. Las bases de microdatos que facilitan la realización del estudio corresponden a la Base de Trabajo (BT) y la Base Usuaría Ampliada (BUA). Por una cuestión de dominio de estimación y para evitar problemas de estacionalidad se consideró, para cada uno de los años, solo la primera onda con relevamiento en el mes de mayo. Estas presentaciones permiten la armonización de sus formatos posibilitando investigar múltiples interrelaciones en el análisis de las variables laborales y sus factores.

II.2. DEFINICIONES CONCEPTUALES.

II.2.1. CENTROS URBANOS.

El análisis de perfiles de trabajadores en condiciones de informalidad laboral se concentra en los grandes centros urbanos de Argentina: Ciudad Autónoma de Buenos Aires, Partidos del Gran Buenos Aires, Gran Rosario, Gran Córdoba, Gran La Plata, Gran Mendoza y Gran Tucumán¹¹.

El criterio técnico propuesto por INDEC (1998) define por aglomeración urbana o localidad compuesta a una región urbanizada que se extiende mediante edificios o calles que conectan en forma continua distintas unidades administrativas. Este principio físico involucra necesariamente a dos o más localidades, comunas, municipios, departamentos o partidos correspondientes a una o más provincias.

¹¹ Las Tablas A.3 a A.5 (g) del anexo “A. Materiales” muestra información referida a estos aglomerados urbanos en estudio.

La estructura de estos conglomerados, en general, distingue dos partes definidas: Un núcleo y una periferia. La primera está compuesta por una ciudad central, de mayor densidad poblacional, donde predominan las actividades que son fuente de empleo para la comunidad. La segunda se compone por ciudades secundarias, de menor densidad, denominadas ciudades satélites o ciudades dormitorio que se conectan mediante vías de transporte directo.

La elección de estos grandes centros urbanos tiene base en criterios concretos que resultan apropiados para mejorar el grado de especificidad en la investigación del fenómeno en estudio. En primer lugar, la población estimada en este subgrupo representa alrededor del 74% de la población contenida por la EPH y aproximadamente el 90% de las aglomeraciones mayores de quinientos mil habitantes que la conforman¹². En segundo lugar, la distribución espacial comprende a cuatro de las seis regiones estadísticas propuestas por INDEC¹³, lo que permite tener cierto nivel de representación geográfica del país. En tercer lugar, debido a que han sido las siete principales aglomeraciones urbanas durante todo el desarrollo de Argentina¹⁴, la evolución de estas ciudades, presenta un componente histórico-cultural muy importante para comprender la realidad del mercado laboral argentino. Por último, en general, es en las principales urbes en el mundo el lugar donde predomina mayoritariamente la relación laboral de producción asalariada que es parte central del objeto de estudio. Asimismo, esta categoría ocupacional se presenta en forma más homogénea debido a que allí se radican las empresas de mayor tamaño. La homogeneidad en la población objetivo favorece en gran medida las estimaciones realizadas mediante las técnicas estadísticas propuestas en esta tesis.

¹² Las Tablas A.4 del anexo “A. Materiales” indica el porcentaje de la cobertura que representan los aglomerados urbanos en el año 2020.

¹³ Región de Gran Buenos Aires (Ciudad Autónoma de Buenos Aires, Partidos de Buenos Aires), Región Pampeana (Gran La Plata, Gran Rosario, Gran Córdoba), Región de Cuyo (Gran Mendoza) y Región Noreste Argentino o NOA (Gran Tucumán).

¹⁴ Las Tablas A.6 del anexo “A. Materiales” brinda información referida a la evolución de los aglomerados urbanos.

II.2.2. TRABAJADORES ASALARIADOS.

En la EPH puntual la condición de actividad de un trabajador ocupado se sostiene sobre la definición histórica de trabajo entendido como “*la actividad laboral general que deriva en la producción de bienes o servicios que tengan valor económico en el mercado*” (Elizalde, Pok, Botta y Villarreal, 1974). En este sentido, este concepto se materializa considerando como ocupados a aquellas personas que hayan trabajado al menos una hora durante la semana de referencia en forma remunerada (mediante dinero o mercadería), que hayan trabajado en forma habitual al menos 15 horas semanales sin pago o que, a pesar de no haber trabajado en la semana de referencia (por vacaciones, licencia o suspensión), mantienen su empleo.

Por su parte, la relación de producción cuya modalidad es el trabajo asalariado se define a partir del empleo en relación de dependencia en cualquiera de sus dimensiones conceptuales. Es decir, comprende el trabajo realizado como empleado u obrero para un patrón, empresa o institución considerando distintos tipos de percepción de pago, temporalidad, nivel de formalidad o forma de contratación. Esto implica que la persona aporta su trabajo personal pero las condiciones organizativas y el capital (instalaciones o maquinaria) le son determinadas, y da lugar a subcategorías de trabajadores asalariados con pago o ad honorem, permanentes, temporales o estacionales, registrados, en negro o en forma encubierta.

II.2.3. EMPLEO ASALARIADO INFORMAL.

La informalidad laboral, en tanto definición conceptual, puede abordarse desde las distintas perspectivas teóricas del mercado de trabajo. Las corrientes que en América Latina han realizado los mayores aportes al estudio de este tópico, corresponden al enfoque legalista o neoliberal, el enfoque estructuralista o neomarxista y los dualistas.

Sin embargo, la presencia de múltiples terminologías sobre un mismo concepto derivó en distintas interpretaciones del problema, agudizando la distancia existente entre estas definiciones y sus indicadores estadísticos. Esto llevó a que el estudio se orientara con mayor énfasis en el método de cálculo de los índices en detrimento de su contenido teórico. Para comienzo del siglo XXI, desde Organización Internacional del Trabajo (OIT), se propusieron las nuevas definiciones de informalidad que fueron ganando hegemonía en estos últimos años principalmente por la operatividad de su medición a partir de datos oficiales.

El criterio de la OIT para determinar el empleo asalariado informal tiene base en lo establecido en la decimoséptima Conferencia Internacional de Estadígrafos del Trabajo (CIET), realizada en diciembre del año 2003 en Ginebra. Esta perspectiva, incluye a todos los empleos asalariados que se caracterizan por una relación laboral desprovista de protección social y legislativa relacionada con el empleo.

En la práctica, con el fin de lograr cierto grado de comparabilidad a nivel internacional, es común que la dimensión de la informalidad en los trabajadores asalariados se reduzca a una única pregunta relacionada al descuento jubilatorio sobre sus ingresos salariales. Es decir, se considera como asalariado no registrado a aquellas personas que, trabajando en relación de dependencia, declaran que no se le ha realizado el descuento jubilatorio en la remuneración de su empleo principal.

Las bases usuarias disponibles en parte del periodo de referencia no contienen una variable de registro estrictamente equivalente¹⁵ que permita realizar el cálculo bajo este único criterio.

¹⁵ Esta variable si está disponibles para el bloque correspondiente a la población desocupada. En este módulo se indaga a la persona sobre la presencia de descuento jubilatorio en su último empleo.

No obstante, la información de las encuestas de hogares, dispone de un módulo de variables que captura los beneficios sociales con los que cuentan los trabajadores asalariados en su ocupación principal. Esta batería de atributos aporta mayor información respecto de la condición laboral del trabajador y posibilita la construcción de un indicador que materialice en forma más precisa el problema en cuestión. En este sentido, permite cuantificar su medición en forma más consistente y controlada.

En conclusión, en este estudio se considera bajo la condición de trabajador asalariado informal a los empleados que en su ocupación principal, desarrollada en relación de dependencia, han sido denegados en forma total o parcial de alguno de los siguientes derechos y/o beneficios laborales: indemnización por despido, vacaciones, aguinaldo, seguro de trabajo, obra social o descuento jubilatorio¹⁶.

II.3. PSEUDO PANEL Y VARIABLES ASOCIADAS (POBLACION Y UN. ANALISIS).

II.3.1. COHORTE POBLACIONAL.

El concepto de cohorte poblacional vincula a un grupo cualquiera de personas que comparten un mismo atributo o han experimentado en un periodo de tiempo algún acontecimiento fijo que los define como tal. Por su parte, es factible que este hecho u atributo sea compuesto por múltiples características que pueden ser muy diversas como el año de nacimiento, el año de graduación escolar, la pertenencia a una actividad profesional, haberse casado o haber publicado un libro, entre otras. En este sentido, las cohortes utilizadas en este trabajo se definen a partir de trabajadores con pertenencia fija a un grupo que pueden identificarse en sucesivas encuestas de hogares.

¹⁶ En cada caso se consideró que estos beneficios sean acordes al tipo de empleo y duración contractual establecida: permanentes, temporarios (a plazo, tarea u obra), changas, inestables.

En términos concretos, para estudiar esta problemática, se definieron cohortes de trabajadores asalariados que compartían simultáneamente dos características comunes: el año de nacimiento y el género declarado. Así, por ejemplo, aquellos trabajadores asalariados que en la base de datos correspondiente al año de inicio del estudio (1990) tenían 20 años de edad, y a su vez, indicaron ser varones pasaron a constituir una cohorte específica que fue identificada, en este caso, como la cohorte “v20”. Las observaciones que se conglomeraron en la cohorte “v20”, tal como se indicó, no solo compartían el género, sino que también comparten el año de nacimiento (1970). En otras palabras, pero con igual validez, podemos indicar que esta cohorte es equivalente a la generación de trabajadores varones nacidos en el año 1970.

Es, entonces, en función de estas dos variables-factores que esta cohorte puede ser identificada, y vinculada, en una base de datos perteneciente al mismo operativo estadístico de corte transversal, pero con información recolectada en una fecha calendario posterior. Por lo tanto, continuando el ejemplo previo, esta cohorte de individuos podrá ser identificada en una base de datos del año 1993 a partir de la agregación de las observaciones que, siendo de género varón, tengan la edad de 23 años. Mientras que, en el año 2003 corresponderá al grupo varones con 33 años de edad. Es decir, la edad de los trabajadores en la cohorte se modifican pero lo que se mantiene fijo es su año de nacimiento, es decir, su generación.

En la construcción de estas cohortes que se definen como subgrupos de la población se deben seleccionar variables-factores que deben ser o suponerse invariantes en el tiempo (por ejemplo, año de nacimiento, género, etnia). A su vez, garantizar que las cohortes se construyan sobre una población y criterios estables que permitan que sus perfiles no cambien bruscamente en el tiempo.

Este concepto es fundamental para comprender la importancia de la problemática metodológica presentada en el punto I.3, como así también, los criterios en la elección de las técnicas estadísticas utilizadas debido a que los estudios de individuos seguidos en mediciones a través del tiempo, es decir estudios de cohorte, constituyen un elemento imprescindible para el análisis longitudinal.

II.3.2. BASES DE DATOS DE TRABAJO.

Inicialmente, el procesamiento y análisis de los datos se realiza sobre las bases usuarias de corte transversal descriptas en el apartado II.1.3. Estas proveen la información de los “N” individuos relevados por el operativo, con “P” mediciones de sus atributos de interés que se representan en un vector para cada observación por “ y_i ” para $i = 1, \dots, N$. Este procedimiento, se realiza para cada uno de los $t = 1, \dots, T$ momentos específicos que componen el período en estudio, lo que en símbolos se expresa por y_{it} .

En términos formales, para la construcción de cada una de las C cohortes se agregan aquellos individuos “i”, que según las variables-factores seleccionadas previamente, pertenecen a la cohorte “c”, para $c = 1, \dots, C$. Por lo tanto, se reemplazan los valores de las observaciones individuales “i”, de la base de datos original, con los valores medios de los subgrupos de la población constituyendo una nueva base de datos.

En este sentido, se confecciona una nueva base de datos de trabajo que presenta información de las “C” distintas cohortes de trabajadores asalariados, en las P mediciones de sus atributos de interés en cada uno de los $t = 1, \dots, T$ momentos específicos que componen el período en estudio,

lo que en símbolos se expresa como $\bar{y}_{ct} = \frac{1}{n_c} \sum_{i=1}^{n_c} y_{it}$.

En consecuencia, mientras que la base de datos usuaria, está constituida por las unidades de observación correspondiente al conjunto de los “N” individuos y la medición de sus atributos en el momento “t”. La base de datos de trabajo se construye sustituyendo las unidades de observación por el conjunto de las $C = \frac{N}{n_c}$ cohortes de trabajadores asalariados y los atributos corresponden a la media estimada, para cada subgrupo, de las variables de la base de datos original¹⁷.

La elaboración de la base de datos de trabajo concluye con el proceso de agrupamiento o “data matching” de los “datasets” de cohortes de trabajadores construidos en cada uno de las mediciones realizadas en los periodos de análisis. Esto es posible, dado que la aparición de las cohortes de trabajadores se puede identificar en las repetidas encuestas transversales, es decir, en las bases de datos que pertenecen a cada uno de los distintos momentos “t” del periodo de análisis.

Este último procedimiento posibilita incorporar la dimensión temporal en el conjunto de microdatos¹⁸ y permite, además, identificar los paneles de observaciones. Los paneles corresponden a las cohortes de trabajadores de un mismo estudio transversal pero en varios períodos de tiempo que, al no estar conformados estrictamente por los mismos individuos sino por un proxy que deriva de la estimación media poblacional, reciben el nombre de “pseudo-paneles”. Luego, cada pseudo-panel se sustenta de la información aportada por su cohorte en cada uno de los años entre 1990 y 2003 siempre que la misma se mantenga en el intervalo etario

¹⁷ Tenemos que $C = 1, \dots, C$. son las cohortes de trabajadores definidas y n_c es el número de observaciones dentro de la cohorte “c”, que por cuestiones de simplificación en la demostración, supondremos que las n_c son iguales.

¹⁸ Por conjunto de microdatos se hace referencia al agregado conformado por el total bases de microdatos de EPH.

comprendido entre los 18 y 60 años. Así, por ejemplo, algunas cohortes se incorporan a su pseudo-panel en el año calendario en que esta generación cumple los 18 años desde su nacimiento, mientras que, otras salen del conjunto de microdatos al superar los 60 años de edad.

En conclusión, en este trabajo, la población queda definida como las cohortes de trabajadores en condición de informalidad laboral en los grandes centros urbanos de Argentina en el período 1990-2003.

II.3.3. VARIABLES.

Los atributos o variables de las cohortes de trabajadores utilizadas hacen referencia a características sociodemográficas y económicas que son de interés a los fines de nuestro estudio debido a que contribuyen a la comprensión de la dinámica del empleo informal en el mercado de trabajo argentino¹⁹.

II.3.3.1. INFORMALIDAD LABORAL.

La variable indica el nivel o *tasa media de informalidad* en la ocupación principal en cada una de las cohortes de trabajadores asalariados. En otras palabras, este atributo refleja el promedio de trabajadores en relación de dependencia que han sido denegados en forma total o parcial del derecho a indemnización por despido, vacaciones, aguinaldo, seguro de trabajo, obra social o descuento jubilatorio en cada unidad estadística.

II.3.3.2. SOCIODEMOGRAFICAS.

El nivel medio de educación alcanzada por las cohortes de trabajadores asalariados nos brinda información del grado de capacitación oficial adquirida al momento de participar en el mercado

¹⁹ La tabla A.7 del Anexo II “A. Materiales” muestra el diccionario de variables en la base de datos de trabajo utilizadas en el análisis.

de trabajo lo que, empíricamente, es considerado un factor determinante para su participación en el mercado formal (Formichella, M. M. & Rojas, M., 2008).

A su vez, los trabajadores con bajos niveles de instrucción suelen presentar mayor predisposición al trabajo cuentapropista cuando se contrae la demanda de trabajo debido a la mayor selectividad por parte de los empleadores y amplían su contribución al trabajo asalariado en épocas de mayor expansión de la demanda laboral.

En este sentido, se definieron dos niveles educativos de análisis: las proporciones de trabajadores que han finalizado al menos su educación primaria básica y las proporciones de trabajadores que han completado, a lo sumo, su formación educativa secundaria.

La variable *primario completo* indica la tasa o proporción de empleados que no han finalizado el trayecto de educación primaria en la cohorte (siete años en instituciones educativas oficiales). Entre los trabajadores que no han finalizado la educación básica primaria se encuentran aquellos que no han tenido instrucción, que han realizado solo educación inicial o que han realizado parcialmente sus estudios primarios. La educación básica en los trabajadores forma parte de los fundamentos de la empleabilidad, por lo que este último grupo, representa un sector con alto grado de vulnerabilidad para la obtención de trabajos decentes (CIET, 2002).

La variable *secundario completo* indica la tasa o proporción de empleados que no han finalizado el trayecto de educación secundaria en la cohorte (catorce años en instituciones educativas oficiales). Entre los trabajadores que no han finalizado la educación secundaria se encuentran aquellos que no han tenido ninguna instrucción, que han realizado solo educación inicial, que han cursado la educación básica primaria y que han cursado parcialmente sus estudios secundarios. Este último grupo encuentra reducidas sus posibilidades de acceso al mercado formal debido a

que, entre otros factores, al no completar la totalidad de la educación media en Argentina no poseen un título que acredite sus conocimientos (Berg, I., 1970) ²⁰.

Por su parte, la variable *edad* representa el tiempo que han vivido los trabajadores que componen cada cohorte, en cada uno de los momentos en que realiza mediciones en el estudio.

II.3.5. RAMA DE ACTIVIDAD ECONÓMICA.

En relación a las ramas de actividad, en la que los trabajadores desempeñan su tarea principal, fueron consideradas: *Industria, Comercio y Construcción*. La nomenclatura utilizada corresponde a la Clasificación de Actividades Económicas para Encuestas Sociodemográficas (CAES) homogeneizando los cambios de codificación de bases a partir de la Clasificación Industrial Internacional Uniforme (CIU) revisión segunda y tercera. En cada caso se tomó en cuenta la proporción de trabajadores en la cohorte cuya ocupación principal se desarrollaba en alguno de los sectores económicos mencionados.

II.3.6. INTENSIDAD OCUPACIONAL.

La variable intensidad ocupacional representada por la *Proporción de subocupación horaria en la cohorte* mide la cantidad de horas que un trabajador dedicó en la semana a sus actividades laborales. Tanto el exceso como la insuficiencia en las horas trabajadas son, en términos generales, un reflejo indirecto de precariedad laboral. En el primer caso debido a la incapacidad para satisfacer necesidades básicas y en el segundo se lo relaciona con niveles de informalidad. La intensidad laboral menor a 35 horas semanales se designa por Subocupación por insuficiencia horaria, mientras que, una dedicación mayor a 45 horas se la denomina Sobreocupación. En

²⁰ Este fenómeno se denomina “*credencialismo*” y su teoría concibe a la educación como una credencial necesaria para el acceso al mercado de trabajo formal.

consecuencia, ocupados plenos, son aquellos trabajadores cuyas ocupaciones laborales demandaron entre 35 hasta 45 horas semanales.

Otra variable considerada fue la antigüedad en el puesto de trabajo medido como la *proporción que permanecieron más de un año en su puesto de trabajo* en la cohorte.

II.3.7. EMPRESA O ESTABLECIMIENTO LABORAL.

El tamaño de la empresa indica la dimensión que posee la organización, en término de la cantidad de empleados que contiene. Esta variable da un reflejo del registro de la ocupación. El no registro, tanto de la actividad como del empleo, suele presentar mayores niveles en empresas pequeñas o microemprendimientos. La variable considerada es la *proporción de trabajadores en establecimientos de menos de 5 empleados* en la cohorte.

II.3.8. TEMPORAL.

La variable *Año* indica a cada uno de los momentos (o “*timings*”) en que se realizan mediciones sobre las cohortes. En cada caso, corresponde a la primera onda (mayo) del año calendario que señala su etiqueta.

La *Etapa* indica los puntos de corte de la serie, es decir, en donde se admite la posibilidad de cambio estructural. El nivel 1 incluye el intervalo de años entre 1990 y 1994; el nivel 2 abarca el periodo entre 1995 y 1998; y el nivel 3 se conforma con los años comprendidos entre 1999 y 2003.

II.3.9. INGRESO MONETARIO.

El *ingreso monetario recibido en la ocupación principal por una hora de trabajo* es una medida que nos permite comparar el poder adquisitivo percibido en la actividad principal, establecida

como aquella que demanda al trabajador mayor cantidad de horas, durante una misma unidad de tiempo (hora).

La hiperinflación monetaria sufrida por Argentina a inicio del periodo provoco una fuerte inestabilidad en el valor de la moneda. A su vez, el cambio establecido en el sistema monetario argentino mediante la Ley 23.938 de Convertibilidad del Austral, y el periodo de adaptación que mantuvo fluctuaciones en la moneda, dificultan la comparación del poder adquisitivo del salario percibido por los trabajadores en su ocupación principal. En consecuencia, se calculó el *ingreso promedio por hora estandarizado* por las cohortes de trabajadores.

PARTE III. ANALISIS EXPLORATORIO DE LAS VARIABLES.

En el tercer capítulo se realiza el análisis exploratorio de las variables a partir de gráficos y medidas descriptivas multivariadas.

En la primera sección (III.1), se estudia la distribución de las variables en el conjunto de datos sin valorar la existencia de paneles o la presencia de estructura temporal en el mismo. Esto permite focalizar el análisis en el aporte que cada cohorte realiza en la distribución de las variables. Inicialmente, se describen las principales características de la variable informalidad laboral y su relación con el resto de las variables en estudio. Luego, la distribución del empleo informal es particionada según el año de relevamiento para identificar e interpretar los atípicos transversales. Finalmente, se evalúa la interrelación existente entre las variables más relevantes para explicar el comportamiento de mercado de trabajo argentino.

En la segunda sección (III.2), se observa el comportamiento de los valores promedios anuales de las variables definidas en este trabajo a través del tiempo. El análisis se realiza desagregando la evolución de los valores medios anual según aglomerado urbano, género y las etapas de quiebre temporal descriptas con anterioridad en el apartado II.3.8.

III.1. ANÁLISIS DESCRIPTIVO: DATOS AGRUPADOS.

En esta primera sección se estudia la distribución del conjunto “N” en función de medidas descriptivas, que son calculadas sin valorar la existencia de paneles o la presencia de estructura temporal en el mismo.

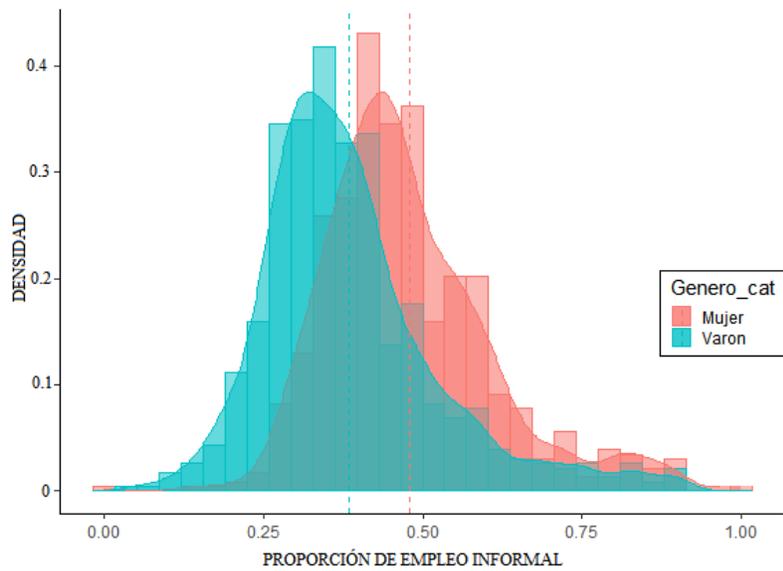
Este primer análisis de datos agrupados, tal como se mencionó, no valora la dimensión temporal contenida en el conjunto de datos pero permite identificar el aporte individual que realiza la cohorte “c” en las variables²¹. En otros términos, en esta instancia exploratoria, se evalúa el conjunto de datos como un conjunto de información transversal omitiendo el factor temporal.

Partiendo del criterio mencionado, inicialmente, se describen las principales características de la variable informalidad laboral y las relaciones existentes con las variables de interés descriptas en la sección II.3.3.

El histograma (Figura III.1) permite visualizar que la distribución es relativamente simétrica con presencia de algunas cohortes que admiten niveles de empleo informal altos en relación a los niveles medio registrados en cada género. En las cohortes de varones la concentración de las observaciones en la distribución, y su promedio, se ubican a valores menores a los asumidos en las cohortes de trabajadoras. La asimetría positiva exhibida en su forma (0.91) es confirmada en la prueba presentada en la Tabla B.1 del Anexo III “*Análisis Exploratorio*”.

²¹ Al no considerar existencia de paneles, se puede identificar en la distribución los valores asumidos por cada cohorte en cada una de las variables en cada momento específico de tiempo como si se tratase observaciones independientes.

FIGURA III.1. Histograma: Empleo informal según género.



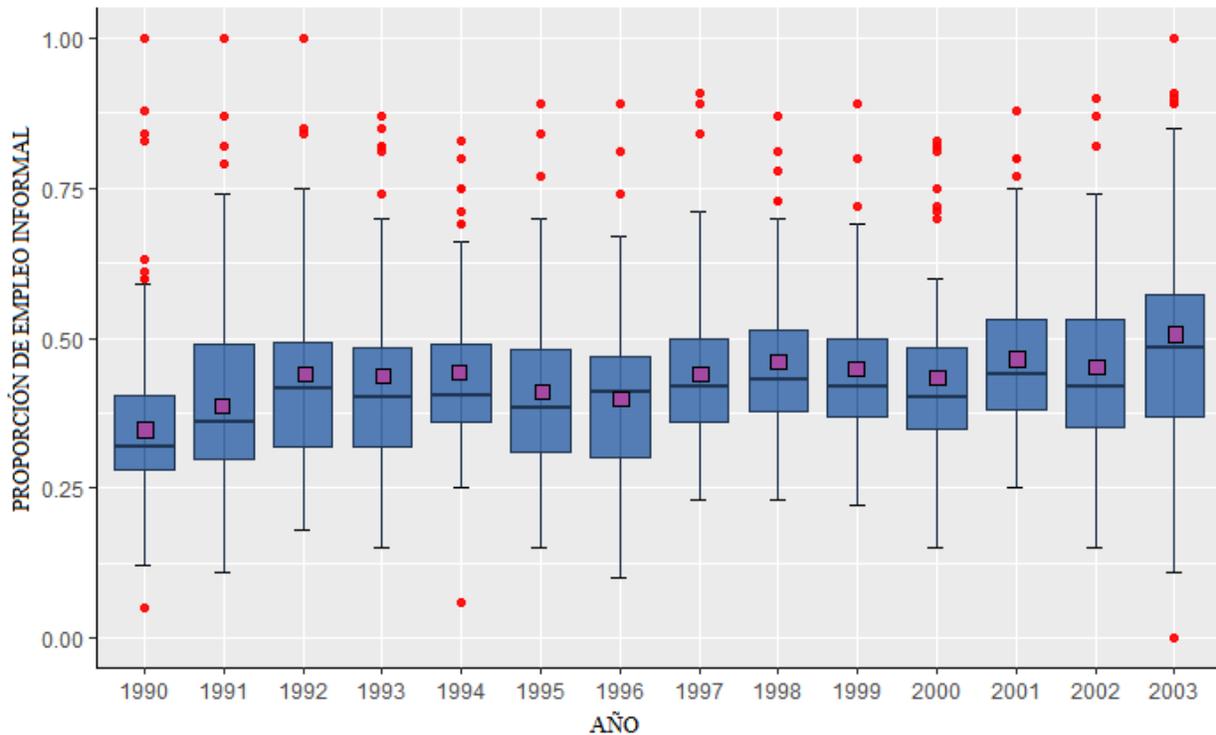
NOTA: La figura muestra las densidades de la proporción del empleo informal de cada género.

Una primera incorporación de la dimensión temporal es posible si se considera la distribución de la variable del conjunto de cohortes en cada uno de los años en estudio. Esto es presentado gráficamente en la Figura III.2. En los diagramas de caja y brazo de los años en estudio, la variable presentó una marcada asimetría positiva, en donde se observa que el valor mediano de la informalidad laboral representado con la barra en el centro de la caja fue, en general, menor al valor promedio de la variable en casi todos los años indicado con un recuadro lila. La excepción se observa en el año 1996, en donde la informalidad promedio se ubicó por debajo del valor que cubre al cincuenta por ciento de las cohortes de menor empleo informal.

Por su parte, la dimensión de la caja permite dar una idea de la heterogeneidad entre las cohortes ubicadas en el 50 por ciento central de la distribución. Es posible distinguir, que en los años en donde los niveles medios de informalidad presentaron variaciones significativas, se advirtió una

mayor amplitud entre los límites de su distribución central. Una potencial hipótesis a corroborar podría ser si en años de crisis económicas las cohortes de trabajadores tienden a presentar mayor heterogeneidad en sus niveles de informalidad.

FIGURA III.2. Diagrama de caja y brazos: Distribución de empleo informal en cohortes según año.



NOTA: La figura muestra las distribución de empleo informal en las cohortes de trabajadores asalariados en cada uno de los años en estudio. Además se indican valores atípicos (puntos rojos), mediana (barra horizontal) y media (rectángulo lila).

Una posible explicación de la deformación en la distribución puede deberse a la influencia que ejercen las cohortes con valores extremos sobre el cálculo de la media aritmética. En relación a esto, los puntos rojos²², indican la presencia de atípicos por derecha en todos los años y por izquierda en los años 1990, 1994 y 2003.

²² Los puntos rojos indican cohortes con niveles de informalidad que se alejan del tercer cuartil en más de 1,5 veces el recorrido intercuantílico. Mientras que, la figura III.7 del Anexo III se indica las cohortes

La Figura B.1 y la Tabla B.4 del Anexo III permiten identificar la ubicación de las cohortes en la distribución, la intensidad de su proporción de empleo informal, como así también, aquellas de comportamiento atípico en un año particular²³. Un dato relevante es que valores atípicos por izquierda observados, es decir, aquellos que presentaron niveles de informalidad extremadamente bajos correspondieron en todos los casos a cohortes de trabajadores varones.

En cuanto a los valores atípicos por derecha, durante todo el período se registraron 51 casos. Entre las características más notorias (Tablas B.3 y B.4) se destaca que el 72.5 por ciento correspondieron a edades de ingreso al mercado de trabajo comprendidas entre los 18 y 21 años, mientras que, el 25.5 por ciento contempló cohortes con edades avanzadas en el mercado de trabajo, mayores a 60 años de edad. Asimismo, el 68.6 por ciento fueron generaciones de mujeres. Es decir, en aproximadamente dos de cada tres casos se trató de cohortes femeninas y en uno de cada tres hace referencia a cohortes de varones.

Por último, las cohortes con valores de alta informalidad se repitieron principalmente entre los años 2000 y 2003, en donde las cohortes de mujeres con edades de ingreso en el mercado laboral (“m06”, “m07”, “m08”) aparecen en forma reiterada y sucesiva con valores de informalidad considerablemente elevados. Esto estaría indicando una prolongación del tiempo medio de ingreso en el mercado formal de trabajo en estos años.

Las principales relaciones entre el empleo informal y las demás características sociodemográficas y económicas son presentadas a continuación mediante el cálculo de los estadísticos de correlación y los gráficos de dispersión conjunta.

²³ En la Figura B.1 se indican los “ID” de las observaciones con niveles de empleo informal mínimo y máximo en cada año específico.

La Tabla III.1 muestra los resultados del coeficiente de Pearson para medir la correlación lineal entre la proporción de informalidad y las demás características analizadas en las cohortes de trabajadores asalariados de los grandes centros urbanos de Argentina en el período 1990 a 2003. La informalidad laboral se correlaciona en forma positiva con los niveles de formación educativa incompleta (primaria y secundaria), con la baja intensidad horaria en la jornada laboral, con la proporción de empleo en establecimientos con menos de 5 empleados, con el empleo en actividades del sector comercio y con el transcurso del tiempo. Estas relaciones son estadísticamente significativas a un nivel de significación de 1 por ciento, a excepción de la capacitación en educación secundaria.

Por su parte, los atributos de las cohortes de trabajadores que se correlacionaron en forma negativa con la informalidad fueron la edad de la cohorte (expresada en años), la proporción de empleados con antigüedad mayor a un año en el puesto de trabajo, el ingreso medio percibido por hora en ese trabajo (estandarizado) y proporción de empleos con tareas realizadas en los sectores de la industria y la construcción. Es decir, cuando estas características aumentan los niveles de informalidad en la cohorte disminuyen

La correlación más alta en el caso positivo, se presentó entre la informalidad con el tamaño del establecimiento (0.69) y la intensidad laboral (0.49). Mientras que, en el caso negativo, fue entre la informalidad y la antigüedad en el puesto de trabajo (-0.63). En la Figura III.3 se muestran en forma gráfica las principales relaciones mencionadas, en tanto, la Figura B.2 permite visualizarlas desagregadas según el género de la cohorte.

Tabla III.1.

Correlación entre la la tasa de informalidad y las características de los asalariados en el período 1990-2003. Método "Pearson's product-moment correlation". Bilateral.

Informalidad	Coefficiente	IC Inferior	IC Superior	Estadístico	p-valor
Año	0.17	0.12	0.22	6.3	0.0000
Primario incompleto	0.07	0.02	0.12	2.6	0.0088
Secundario incompleto	0.04	-0.01	0.09	1.5	0.1409
Rama Industrial	-0.25	-0.30	-0.20	-9.6	0.0000
Rama Comercio	0.31	0.26	0.35	11.8	0.0000
Rama Construcción	-0.12	-0.17	-0.07	-4.4	0.0000
Intensidad laboral	0.49	0.45	0.53	20.4	0.0000
Antigüedad laboral	-0.63	-0.66	-0.59	-29.5	0.0000
Tamaño establecimiento	0.69	0.66	0.72	35.0	0.0000
Ingreso por hora	-0.13	-0.18	-0.07	-4.6	0.0000
Ingreso por hora estandarizado	-0.19	-0.24	-0.14	-7.1	0.0000
Edad	-0.28	-0.33	-0.23	-10.6	0.0000

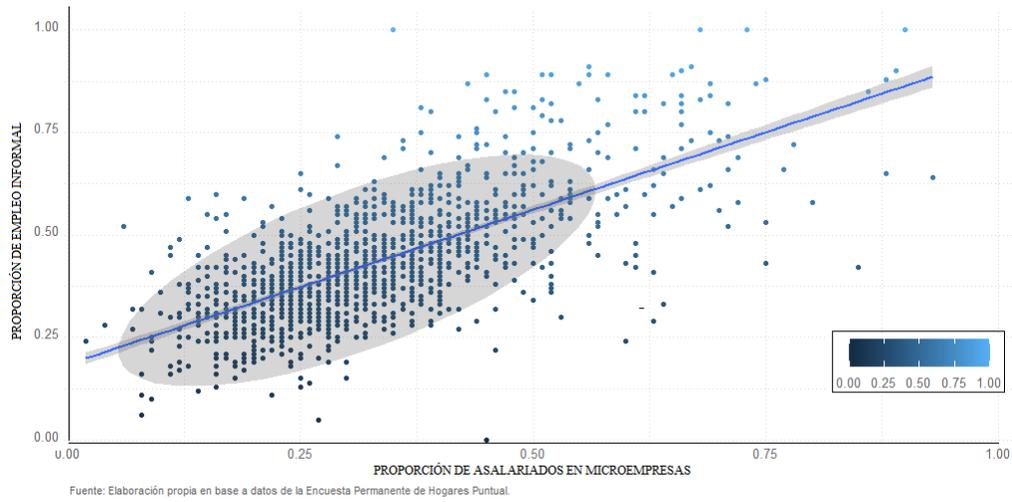
FUENTE: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual..

La distribución conjunta de la informalidad laboral y el tamaño del establecimiento empleador, presenta su mayor concentración cuando los porcentajes de empleo en microempresas son menores al 50 por ciento. Al desagregar esta relación según el género de la cohorte se observa que las proporciones de mujeres en microempresas, en términos medios, son sustancialmente mayores.

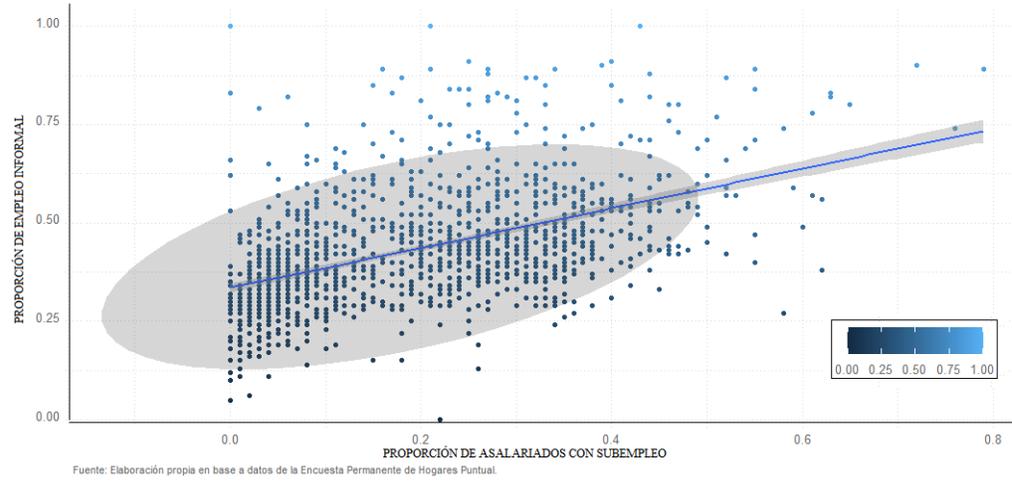
La segregación en las cohortes de trabajadores se reitera también en el caso de la intensidad laboral. Las cohortes de mujeres se aglomeran en proporciones de subocupación con valores mayores al 20 por ciento. Por el contrario, las cohortes de varones se concentran en proporciones de subocupación menores a este valor.

FIGURA III.3. Diagrama de dispersión.

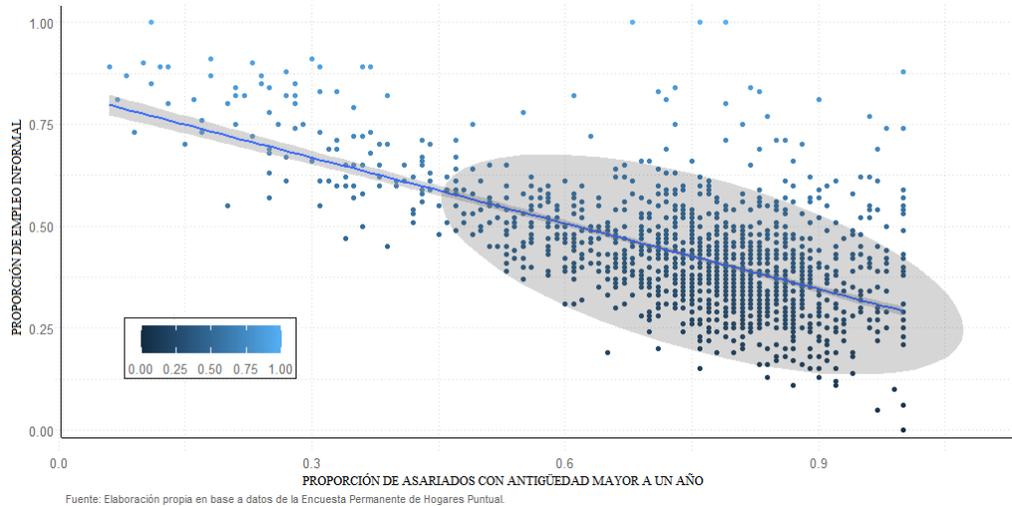
(a) Informalidad Laboral y Tamaño establecimiento.



(b) Informalidad Laboral e Intensidad Horaria.



(c) Informalidad Laboral y Antigüedad en la ocupación.



Sin embargo, esta discrepancia no sucede cuando se analiza la antigüedad laboral. En este caso, la proporción de asalariados con antigüedad mayor a un año en las tareas realizadas en su ocupación principal se distribuye en forma similar para ambos géneros.

En relación a la distribución de las demás variables de interés y su vinculación conjunta, a continuación se analizan las principales medidas descriptivas multivariadas.

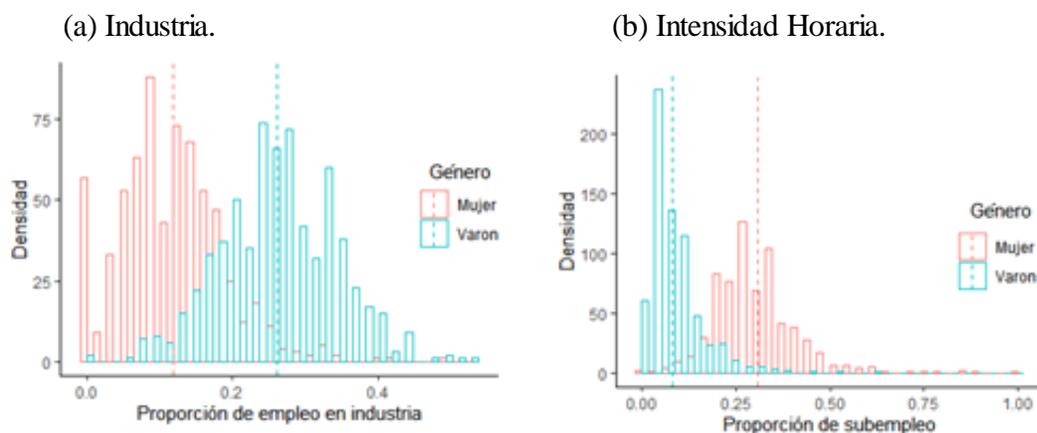
Los coeficientes obtenidos en la Tabla B.1 del Anexo III, indican presencia de asimetría positiva en la distribución de las variables que representan a los ingresos medios percibidos por hora, la participación promedio en actividades de construcción o comercio, la proporción de empleos en microempresas y la baja intensidad horaria. Esto indicaría existencia de algunas cohortes de trabajadores asalariados con valores relativamente altos en el rango de la distribución de estos atributos. Las distribuciones que presentaron marcada asimetría negativa o izquierda corresponden al promedio de cohortes de trabajadores con nivel instrucción primaria incompleta y a la proporción de cohortes con empleos con antigüedad en su ocupación mayor a un año.

La Tabla B.6 del Anexo III, evalúa las aproximaciones a la distribución normal mediante los estadísticos de las principales pruebas de hipótesis. A excepción de la proporción de trabajadores con secundario incompleto en las cohortes, el supuesto de normalidad no se cumple al nivel del 5 por ciento.

En cuanto al comportamiento de las variables desagregadas según el género de la cohorte, los histogramas de la Figura III.4, permiten identificar marcadas diferencias en las distribuciones de la intensidad horaria y la participación en la industria. En este sentido, la proporción de subempleo en las cohortes masculinas es notoriamente más asimétrica que en el caso de las mujeres, lo que indicaría una alta concentración de las cohortes con niveles medios de subempleo

relativamente bajos en relación a las cohortes femeninas. Por su parte, la proporción de empleo en actividades de la rama industrial, en el caso de los asalariados varones, refleja una distribución relativamente simétrica en relación a las asalariadas mujeres.

FIGURA III.4. Histograma: Intesidad Horaria y Empleo en Industria según género.



NOTA: La figura muestra las densidades de la proporción del empleo en industria y la proporción en las cohortes de trabajadores según cada género.

La intensidad de la variación conjunta entre variables se expresa en la Tabla III.2. La matriz de correlación total involucra a todos los atributos de las cohortes de trabajadores que se han analizado previamente. En este sentido, nos permite evaluar simultáneamente la intensidad de la relación lineal existente entre las características sociodemográficas, laborales o económicas que están en estudio en el conjunto de datos.

Tabla III.2.

Correlación entre la la tasa de informalidad y las características de los asalariados en el período 1990-2003. Método "Pearson's product-moment correlation". Bilateral.

	Informalidad	Año	Prim. Incompl.	Sec. Incompl.	Industria	Comercio	Construccion	Intensidad	Antigüedad *	Tam_Estab	Edad	Ing. Hora Estand.
Informalidad	1.00	0.17	0.07	0.04	-0.25	0.31	-0.12	0.49	-0.63	0.69	-0.28	-0.19
Año	0.17	1.00	0.21	0.26	-0.30	0.02	0.01	0.21	-0.07	0.01	0.00	0.52
Prim. Incompl.	0.07	0.21	1.00	0.57	-0.06	0.34	0.01	-0.02	-0.40	-0.09	-0.68	0.03
Sec. Incompl.	0.04	0.26	0.57	1.00	-0.37	0.12	-0.37	0.31	-0.20	-0.05	-0.50	0.15
Industria	-0.25	-0.30	-0.06	-0.37	1.00	0.10	0.46	-0.63	-0.03	-0.38	-0.11	-0.18
Comercio	0.31	0.02	0.34	0.12	0.10	1.00	0.16	-0.11	-0.61	0.17	-0.54	-0.23
Construccion	-0.12	0.01	0.01	-0.37	0.46	0.16	1.00	-0.55	-0.14	-0.26	-0.13	-0.10
Intensidad	0.49	0.21	-0.02	0.31	-0.63	-0.11	-0.55	1.00	-0.15	0.53	0.01	0.08
Antigüedad *	-0.63	-0.07	-0.40	-0.20	-0.03	-0.61	-0.14	-0.15	1.00	-0.38	0.73	0.31
Tam_Estab	0.69	0.01	-0.09	-0.05	-0.38	0.17	-0.26	0.53	-0.38	1.00	-0.05	-0.23
Edad	-0.28	0.00	-0.68	-0.50	-0.11	-0.54	-0.13	0.01	0.73	-0.05	1.00	0.25
Ing. Hora Estand.	-0.19	0.52	0.03	0.15	-0.18	-0.23	-0.10	0.08	0.31	-0.23	0.25	1.00

Fuente: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual.

* Proporción de antigüedad mayor de un año en la cohorte.

Las principales correlaciones negativas que se identificaron vinculan al empleo en la industria con la subocupación horaria (-0.63), las tareas en comercio con la antigüedad laboral (-0.61) y con la edad de los trabajadores (-0.54); y a las actividades en construcción con la subocupación horaria (-0.55). Asimismo, se admite una relación lineal negativa entre los niveles de instrucción incompleta, primaria (-0.68) y secundaria (-0.50), respecto de la edad de la cohorte.

En tanto, las principales correlaciones positivas que se identificaron vinculan el tamaño de la empresa donde se desarrolla la actividad con la baja intensidad laboral (0.53), el transcurso de los años con la distribución del ingreso por hora (0.52) y la antigüedad laboral con la edad de la cohorte (0.73).

III.2. ANÁLISIS DESCRIPTIVO: VARIABLES EN EL TIEMPO.

En esta sección, se describe el comportamiento de las variables definidas en este trabajo a través del tiempo utilizando sus valores promedios anuales en relación con el nivel de informalidad, desagregando el análisis según aglomerado urbano y las etapas de quiebre temporal descriptas con anterioridad en el apartado II.3.8.

III.2.1. TRABAJO ASALARIADO.

En términos históricos, entre el 70 y el 75 por ciento de los trabajadores ocupados en Argentina han sido contenidos en la categoría asalariada. Los últimos datos disponibles en la Encuesta Permanente de Hogares, correspondiente al tercer trimestre de 2020, estiman que esta categoría estaba compuesta con alrededor de 7,6 millones de personas, lo cual es equivalente al 71,5 por ciento de los trabajadores ocupados en el país en ese trimestre. Este porcentaje alcanza el

72,3%²⁴ para el los aglomerados en estudio (decir cuales) representan el 74% del total de trabajadores ocupados (5.640.000 trabajadores)

En el período 1990-2003, la proporción de asalariados en la población ocupada respecto de las demás categorías ocupacionales, se mantuvo en constante ascenso (de 68,7% en 1990, a cerca del 73% al inicio del 200)

El porcentaje de asalariados promedio para toda la serie es del 70.6 por ciento, presentando un valor mínimo de 67.6 (%) en el año 1993 y alcanzando su valor máximo de 73.0 (%) en el año 2000²⁵.

En relación a la informalidad laboral, para esta categoría ocupacional, el menor valor del 35.6% se presentó en el primer año de la serie, mientras que el mayor del 50,6%, en el último año. El nivel medio de la informalidad en los trabajadores asalariados fue de 43,9% (DE: 3,5%) para toda la serie²⁶.

Este grupo exhibió dos importantes aumentos en la informalidad durante el periodo de referencia. Estos saltos en los niveles estimados de la variable modificaron el dígito de la decena en las transiciones entre los años 1990-1991 y 2002-2003.

²⁴ Instituto Nacional de Estadísticas y Censos (2020). Informes técnicos. Vol. 4, nº 231. Serie Trabajo e ingresos. Vol. 4, nº 7. "Mercado de trabajo. Tasas e indicadores socioeconómicos (EPH)". Tercer trimestre de 2020.

²⁵ La Tabla B.7 del Anexo III muestra el porcentaje de personas ocupadas según categorías ocupacionales en los grandes centros urbanos de Argentina desagregado según cada uno de los años en estudio.

²⁶ La tabla B.8 del Anexo III muestra las medidas descriptivas correspondientes a la tasa de informalidad en los grandes centros urbanos de Argentina en el período 1990-2003.

III.2.2. TASA DE DESOCUPACIÓN.

La Tasa de Desocupación²⁷, junto a la Tasa de Actividad (TA) y la Tasa de Empleo (TE), constituyen los indicadores fundamentales para explicar la dinámica del mercado de trabajo. Sin embargo, es la primera de estas la que presenta mayor correlación con la Tasa de Informalidad laboral (TI). Esta hipótesis se apoya en el enfoque de la economía dual, a la que adhiere línea de la CEPAL, siendo una de las principales perspectivas teóricas para explicar el origen del aumento del empleo informal en el país, a partir de la segunda parte de la década 1970.

La relación entre estos dos indicadores²⁸ supone desde este enfoque teórico que la problemática de la informalidad nace como refugio de la desocupación en los países periféricos.

La evolución de la tasa de informalidad al igual que la de desocupación en la población de asalariados muestra una tendencia ascendente. A inicios de la década de 1990, el porcentaje de desocupación entre los asalariados rondaba el 7,4%. Mientras que, a partir del inicio de la segunda etapa (1995), y hasta el final del periodo (2003), este porcentaje se ubicó por encima de los dos dígitos. Pero esta tendencia cambia por períodos. En el período 1990-1994 se da una relación inversa, que se repite en el periodo 1995-2002 con menor intensidad.

Es probable que los niveles de empleo informalidad sean relevante para explicar las variaciones en la desocupación, principalmente en los momentos de shocks económicos recesivos, y a su vez, es factible interpretar que los niveles de desocupación sean de importancia para explicar las variaciones en la informalidad, especialmente en la creación de nuevos empleos en épocas de

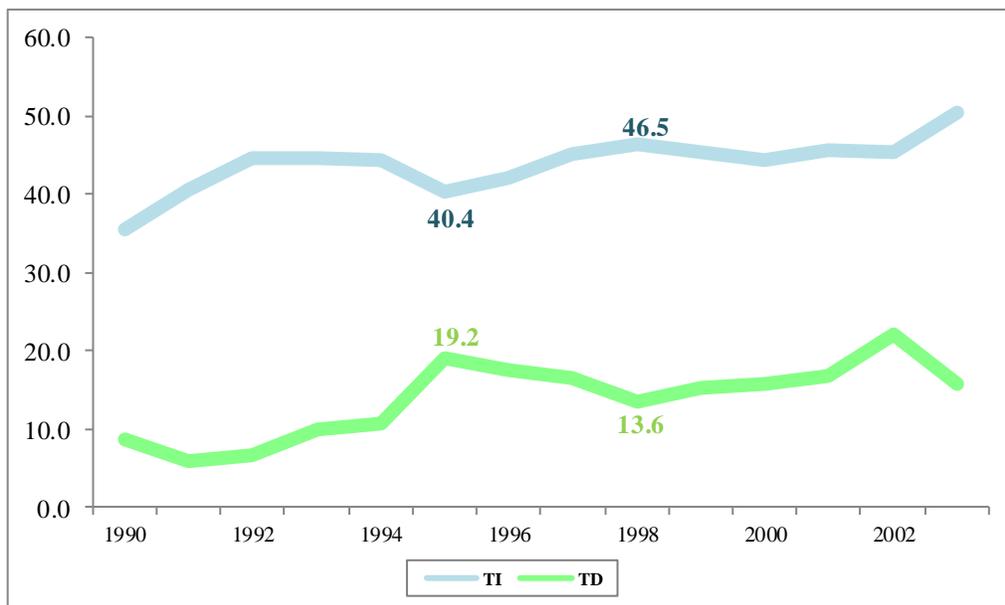
²⁷ Se define como desocupado como aquella persona que actualmente no tiene empleo pero que lo estuvo buscando activamente durante el último mes y se encuentra disponible para empezar a trabajar en caso de conseguir. La Tasa de Desocupación laboral (TD) indica la proporción que representa la Población Desocupada (PD) en relación a la Población Económicamente Activa (PEA).

²⁸ Un trabajo que combina el enfoque de pseudo paneles y series temporales utilizando modelos de Autorregresión Vectorial de Panel (PVAR) para estimar la interrelación entre la informalidad laboral y el desempleo en Argentina es el de Arias, O., & Escudero, W. S. (2007). "Assessing trends in informality in Argentina: A cohorts panel VAR approach." Mimeografía, Banco Mundial y CEDLAS.

reactivación o expansión económica. Sin embargo, la tasa de desocupación debido al pequeño volumen de trabajadores que alcanza dentro de la población económicamente activa no podría explicar la totalidad de la variabilidad del empleo informal o los niveles que este logra en un periodo específico, sino que a lo sumo la participación de un grupo reducido. En este sentido, podría indicar la presencia de una frontera con alta grado de rotación de un conjunto de trabajadores que pasan de la condición de desocupados a la condición de ocupados en condiciones de informalidad, y viceversa.

Esto último, pondría en evidencia la vulnerabilidad del empleo informal. No obstante, no es alcance de este trabajo observar la existencia y las características de la población económicamente activa en la frontera desocupación-informalidad.

FIGURA III.5. Evolución de la tasa de desocupación e informalidad en los grandes centros urbanos de Argentina entre 1990 y 2003.



III.2.3. NIVEL EDUCATIVO ALCANZADO.

En el período considerado, el grado de instrucción de la población ocupada presentó una mejoría en términos cuantitativos. Tanto la participación de trabajadores asalariados con nivel educativo primario incompleto, como la proporción de trabajadores que no concluyeron los estudios secundarios, muestran una sostenida disminución a lo largo de la serie. En trabajadores sin instrucción, esta proporción pasó del 11,5% en el año inicial al 5,5% en el último año de estudio. Esta disminución del 75% de trabajadores con instrucción sólo primaria, es aún más notoria en términos cualitativos ya que sólo el 15% de los trabajadores asalariados cuentan con menos de 13 años de formación educativa. En el caso de trabajadores con secundario incompleto los niveles pasaron del 59,3% al inicio al 44,7% al final del período.

La B.3 del Anexo III permite analizar la proporción media de trabajadores que han completado sus estudios secundarios desagregados por regiones y por los subperíodos o etapas definidas. En este caso, se observa que la formación educativa promedio en la Ciudad de Buenos Aires es mayor que la presentada por los partidos del área metropolitana y por los aglomerados urbanos del interior del país.

El aglomerado Gran Rosario fue el único que presentó una disminución en el nivel promedio de formación secundaria alcanzada en los obreros residentes entre la primera y la última etapa. Esta reducción se vio agudizada principalmente en la segunda etapa, correspondiente a los años 1995 a 1998, donde el porcentaje promedio de trabajadores con educación media cayó en un 11,7%²⁹.

²⁹ Una potencial hipótesis para explicar dicho fenómeno es la migración de mano de obra calificada a otras regiones del país o hacia el exterior debido al alarmante deterioro en las condiciones de vida que se presentaron en este aglomerado principalmente a partir de la segunda etapa en estudio. Los niveles de pobreza se mantuvieron en ascenso llegando a alcanzar al 62,6% de su población residente en el año 2002 (Fuente).

En consecuencia, a partir del deterioro mencionado, este centro urbano pasó del segundo al cuarto lugar en grado de instrucción, siendo desplazado por Gran La Plata y Gran Córdoba³⁰.

Por su parte, los trabajadores asalariados en la Ciudad de Buenos Aires fueron los de mayor instrucción media. Mientras que, en los Partidos de Buenos Aires y el Gran Mendoza, el crecimiento en la formación los obreros, exhibió el mayor crecimiento absoluto y relativo. Pese a esto, estos últimos aglomerados no alcanzaron a superar el nivel promedio de educación secundaria completa del conjunto de los aglomerados en la última etapa del período.

Los centros urbanos Gran Tucumán y Partidos de Buenos Aires se posicionaron por debajo del nivel medio del conjunto en la etapa 1990-1994 y se incorpora a Gran Mendoza si se considera el nivel medio en los años 1999-2003.

III.2.4. RAMAS DE ACTIVIDAD ECONÓMICA (CAES³¹).

En las figuras III.6 a III.8, mediante la altura de las barras, se indica el empleo asalariado promedio que, en los grandes centros urbanos, fue absorbido por actividades de la rama industrial (III.6), comercial (III.7) o (III.8) construcción en cada año observado. Mientras que, las líneas horizontales indican la media aritmética del empleo asalariado contenido en los respectivos sectores económicos según las etapas de estudio (1990-1994, 1995-1998 y 1999-2003).

Por su parte, en las figuras B.3 a B.6 del Anexo B “Análisis Exploratorio” hacen extensivo el análisis desagregando según aglomerado urbano.

³⁰ La discusión respecto al efecto provocado por el régimen de convertibilidad en las condiciones de vida de la población residente de Gran Rosario puede ser profundizada en: Robin, S., & Duran, P. (2005). “Juventud, pobreza y exclusión en el Gran Rosario post devaluación”. En Actas Séptimo Congreso Nacional de Estudios del Trabajo, Buenos Aires.

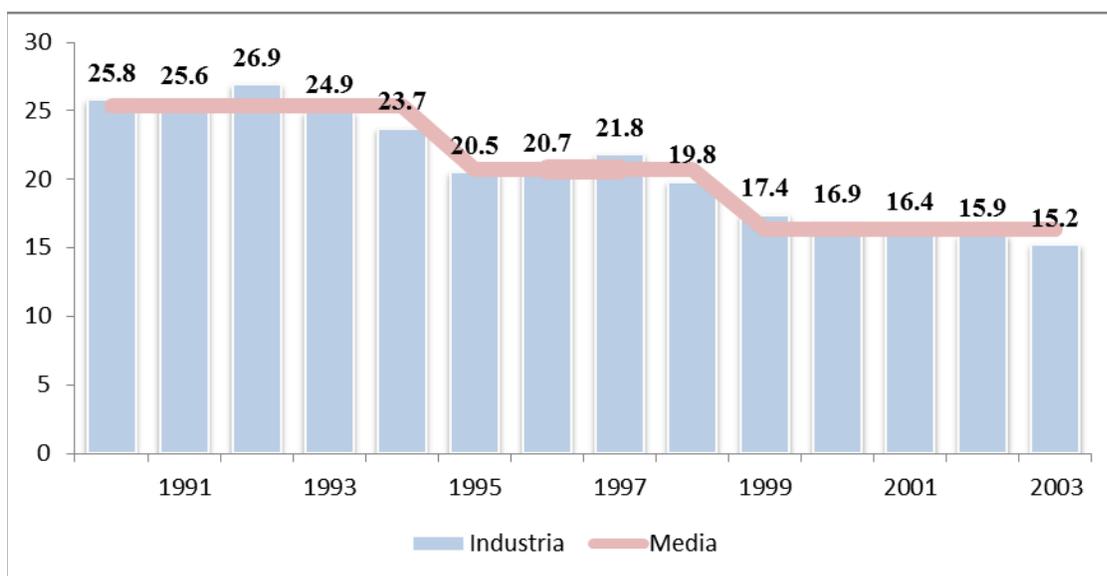
³¹ Tal como se mencionó previamente, para definir las ramas de actividades económicas se utilizó el clasificador de actividades económicas para encuestas sociodemográficas (CAES) realizado por la Coordinación de Clasificadores y Nomenclaturas perteneciente a la Dirección Nacional de Metodología Estadística del Instituto Nacional de Estadística y Censos (INDEC).

III.2.4.1. INDUSTRIA.

La participación del sector industrial para contener el empleo asalariado en los centros urbanos se vio en constante disminución durante todo el período. En el año 1990, el 25,8% de los trabajadores realizaban tareas relacionadas al sector económico industrial. Sin embargo, en la primera onda del año 2003 el porcentaje de obreros en tareas industriales había caído al 15,2%. (Figura III.6).

La figura B.4 del Anexo III desagrega el análisis por localización y periodo. En este caso, se observa que el aglomerado conformado por los Partidos de Buenos Aires es el que empleaba, en forma relativa, la mayor proporción de asalariados en trabajos de la rama industrial y es el centro urbano con mayor disminución en estas actividades.

FIGURA III.6. Proporción de empleo asalariado en actividades del sector industrial. Grandes centros urbanos. Argentina. 1990 a 2003.



NOTA: Las barras indican la proporción de empleo asalariado en actividades del sector industrial en cada año y la línea muestra la proporción promedio en cada etapa.

En los primeros años de la década de 1990 aproximadamente 1 de cada 3 asalariados residentes en los partidos de Buenos Aires se desempeñaba en el sector industrial. En los primeros años de la década del 2000 esa relación se ubicó por debajo de 1 por cada 5 asalariados.

En términos generales, la tendencia se repite en el resto de las localizaciones del país a excepción del aglomerado Gran Rosario, donde se la proporción aumenta al final del período.

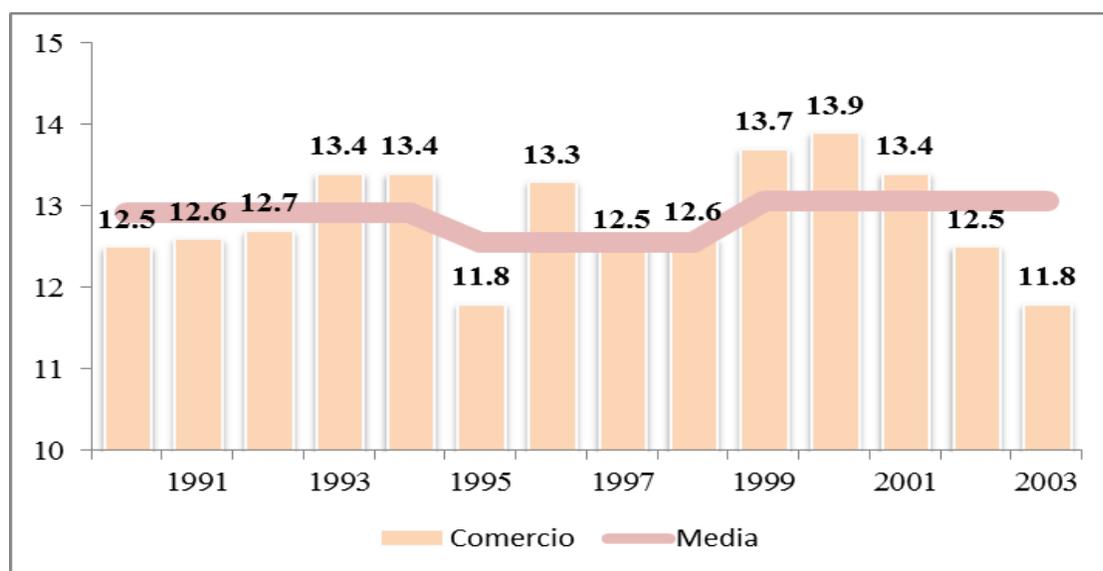
III.2.4.2. COMERCIO.

La participación promedio del comercio como empleador de la mano de obra asalariada presentó un leve aumento (13,1%) durante la etapa en estudio. Sin embargo, es notoria la abrupta caída que exhibe en los últimos años, en donde la variable alcanza su valor mínimo.

Por conglomerado este sector refleja discrepancia en las tendencias expuestas en cada localización. En los centros urbanos de Gran Mendoza, Gran Córdoba y Ciudad de Buenos Aires la tendencia de la variable manifestó disminución en la serie. Mientras que, en Gran La Plata, Gran Rosario, Gran Tucumán y los Partidos de Buenos Aires se mantuvo en aumento.

En los aglomerados de Gran Rosario y Gran Mendoza fue en donde la participación del sector comercio como reclutador de trabajo asalariado tuvo mayor peso relativo. Por su parte, en Gran Tucumán es donde se presentó el mayor crecimiento de la participación de esta rama de actividad en términos absolutos durante el período (figura B.5).

FIGURA III.7. Proporción de empleo asalariado en actividades del sector comercio.
Grandes centros urbanos. Argentina. 1990 a 2003.



NOTA: Las barras indican la proporción de empleo asalariado en actividades del sector comercial en cada año y la línea muestra la proporción promedio en cada etapa.

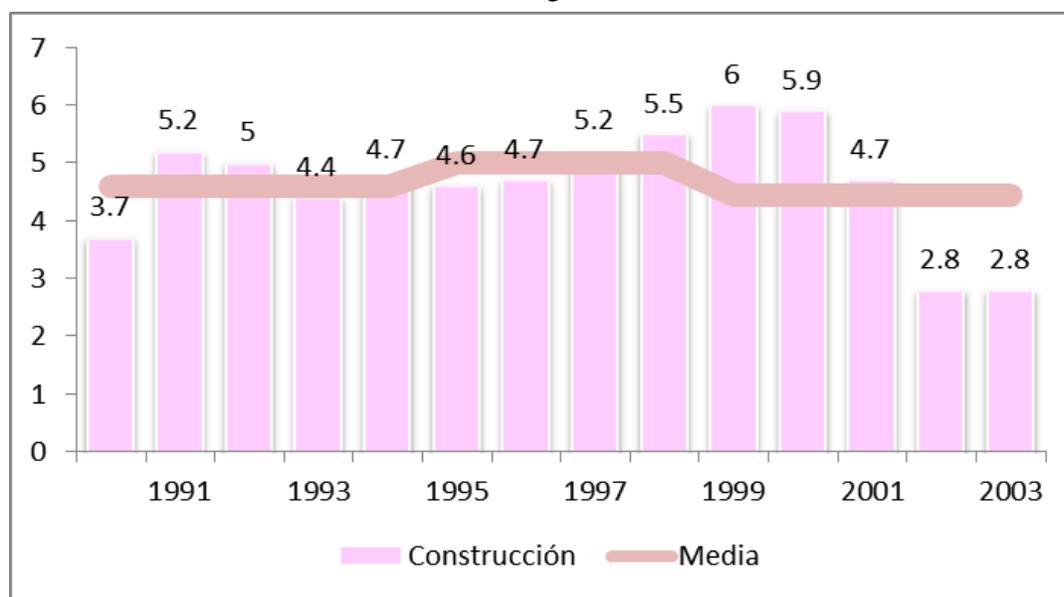
III.2.4.3. CONSTRUCCIÓN.

Para la construcción la tendencia de la serie es poco clara, dado que sus valores exhiben oscilaciones durante todo el período. Una potencial hipótesis para explicar esta variabilidad puede deberse a la sensibilidad existente entre esta variable y la evolución ciclos económicos generales. En este sentido, se diferencian durante este intervalo dos ciclos con picos en los años 1991 y 1999 (Figura III.8).

Por aglomerado se destaca la baja importancia que presenta esta rama de actividad en la ciudad de Buenos Aires con porcentajes inferiores al 2 por ciento en todos los casos.

Por aglomerado se destaca la baja importancia que presenta esta rama de actividad en la ciudad de Buenos Aires con porcentajes inferiores al 2 por ciento en todos los casos.

FIGURA III.8. Proporción de empleo asalariado en actividades de la construcción.
Grandes centros urbanos. Argentina. 1990 a 2003.



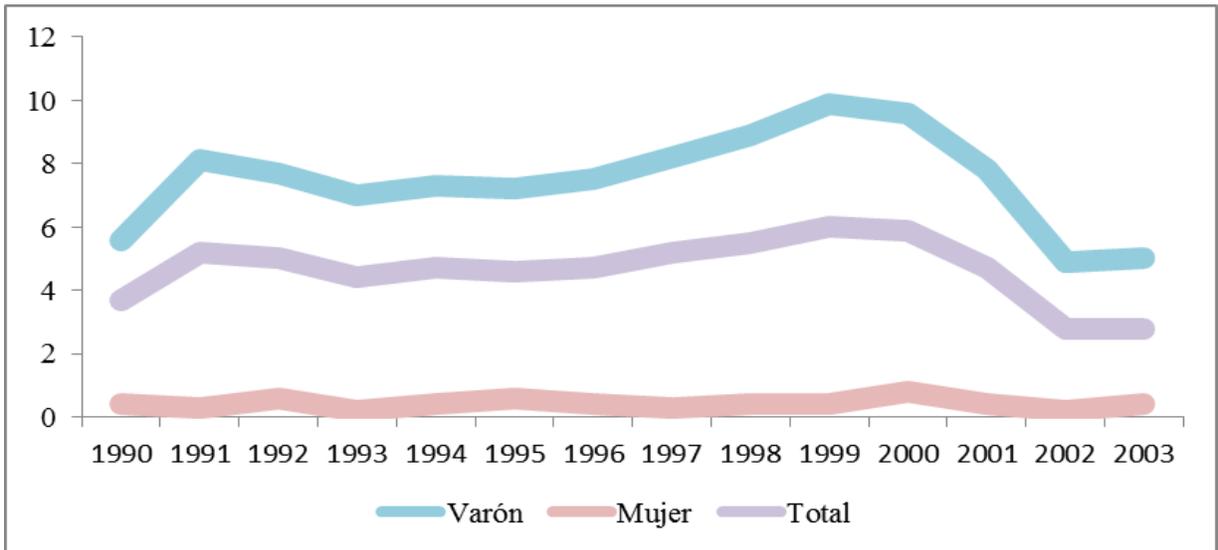
NOTA: Las barras indican la proporción de empleo asalariado en actividades del sector construcción en cada año y la línea muestra la proporción promedio en cada etapa.

Sin embargo, para una correcta interpretación de esta variable, es importante considerar la notoria diferencia de los valores que asume según el género de la cohorte. La figura III.9 muestra dicha segmentación, en donde queda manifiesta la importante heterogeneidad presente.

En consecuencia, el promedio global pierde representatividad para explicar la contribución real de esta rama de actividad como empleadora de la fuerza de trabajo.

En el período de referencia la construcción empleó, en promedio, alrededor del 7,5 por ciento de los obreros varones, alcanzando niveles cercanos al 10 por ciento en el año 1999. Mientras que, en el caso de las obreras mujeres este porcentaje no superó el 1 por ciento en ninguno de los años de la serie.

FIGURA III.9. Proporción de empleo asalariado en actividades de la construcción.
según género. Grandes centros urbanos. Argentina. 1990 a 2003.



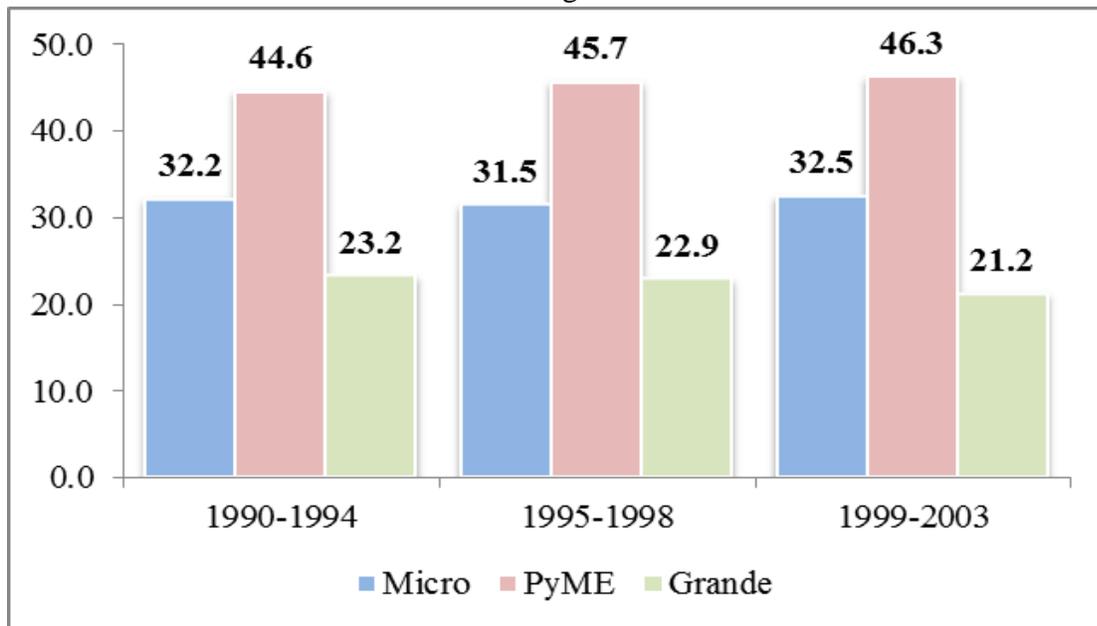
NOTA: Las líneas indican la proporción de empleo asalariado en actividades del sector construcción total y según cada género en cada año.

III.2.5. TAMAÑO DE LOS ESTABLECIMIENTOS.

Las empresas o establecimientos se clasificaron en las categorías de microempresa, pequeñas o medianas empresas y grandes empresas. En relación de la cantidad de empleados que poseen, las Pequeñas y Medianas Empresas (PyMEs) fueron las que contrataron mayor trabajo asalariada durante estos años alcanzando alrededor del 46,3% de los empleados. La Figura III.10 exhibe que en las sucesivas etapas existió un leve ascenso en la participación media de las PyMEs. Este aumento relativo se dio en detrimento de la contribución de las Grandes Empresas a esta variable. Por su parte, el porcentaje de incidencia de las microempresas se mantuvo constante con un nivel de contratación que contuvo a 1 de cada 3 trabajadores asalariados.

Las PyMEs fueron la principal fuente de empleo asalariado en todas las regiones del país (Figura B.7). La única excepción se dio en Gran La Plata, en donde en la primera etapa la contratación por parte de grandes empresas (38,7%) superó al porcentaje de participación de las PyMEs (34,6%).

FIGURA III.10. Proporción de empleo asalariado según tipo de establecimiento. Grandes centros urbanos. Argentina. 1990 a 2003.



NOTA: Las barras indican la proporción de empleo asalariado según el tipo de establecimiento en cada año y la línea muestra la proporción promedio en cada etapa.

Los porcentajes de obreros en microempresas y grandes empresas si mostraron algunas diferencias según localización. La ciudad de Buenos Aires junto a Gran La Plata fueron las conglomeraciones donde mayor presencia tuvo el empleo en empresas de gran tamaño conteniendo en términos medio a poco menos de un tercio de los trabajadores en estudio. En el extremo opuesto se encontró Gran Tucumán donde el porcentaje medio de residentes empleados en microempresas se ubicó próximo al 39 por ciento, mientras que, los ocupados en grandes empresas no alcanzo al 19 por ciento.

PARTE IV. METODOLOGÍA.

En este capítulo se desarrollan analíticamente las metodologías que serán aplicadas para obtener los perfiles sociodemográficos de las cohortes de trabajadores asalariados en los principales centros urbanos de Argentina. En la primera sección (IV.1) se introduce las notaciones básicas de los modelos de panel utilizados comúnmente en econometría, como así también, la conocida discusión entre la estimación bajo inferencia condicional ("efectos fijos") o bajo inferencia incondicional ("efectos aleatorios"). Esta introducción permite describir, en la sección siguiente (IV.2), la extensión a los modelos de pseudo panel, en donde se puntualizan los supuestos que requiere su implementación. Finalmente, en la sección IV.3 se desarrolla en detalle el método longitudinal de clasificación no supervisada utilizado "KmL" (Genolini, C., & Falissard, B., 2010) y la adecuación del criterio de selección óptima de grupos (T. Caliński y J. Harabasz, 1974).

En ambas metodologías, la notación desarrollada fue adecuada al problema objeto de estudio. La extensión o "acople" realizada para incorporar las estimaciones de pseudo panel en el criterio de clasificación longitudinal, permitiendo la combinación de ambas técnicas, constituye un aporte inédito de esta tesis.

IV.1. MODELO DE PANEL.

Los datos de panel hacen referencia a datos que combinan una dimensión temporal con otra transversal. Es decir, un conjunto de individuos que son observados en distintos momentos en el tiempo. El hecho de permitir el seguimiento de las observaciones a través del tiempo proporciona información que posibilita un mejor estudio de las dinámicas de cambio y explicar mejor los fenómenos.

El modelo de panel típico (1) agrega un efecto individual al modelo lineal estándar para capturar el efecto de cada individuo sobre la variable dependiente en el tiempo. Esto se puede expresar mediante la siguiente ecuación:

$$y_{it} = x_{it}\beta + \alpha_i + \varepsilon_{it} \quad (1)$$

$$i = 1, \dots, N. \quad t = 1, \dots, T.$$

y_{it} es la variable de interés para la i -ésima observación en el momento t , x_{it} es el vector de las p variables explicativas, β es el vector de parámetros, α_i es el efecto individual que captura todos los determinantes de la variable de interés que están fijos en el tiempo y ε_{it} es un término de perturbación. El subíndice “ t ” indica al t -ésimo momento del tiempo.

En forma compacta el modelo puede presentarse como sigue:

$$y = X\beta + C\alpha + \varepsilon \quad (2)$$

$$C = I_N \otimes t$$

En donde: y , C y ε son vectores de dimensión NT .

X es una matriz de diseño de dimensión $NT \times P$, β es el vector de P parámetros y $\boldsymbol{\iota}$ es un vector $T \times 1$.

Una primera cuestión a determinar es si la variable aleatoria inobservable α_i está correlacionada con el vector de regresoras x_{it} o si la misma es independiente respecto a las últimas, de lo cual depende el algoritmo de estimación de los parámetros.

En el primer caso, cuando existe correlación, es conveniente realizar la estimación bajo inferencia condicional lo que se denomina modelo de panel de “efectos fijos”. Este modelo es el que presenta menos supuestos respecto a los errores y suele ser el más consistente. En tal caso, α_i asume un valor fijo para cada individuo y es estimado conjuntamente con β mediante el método de mínimos cuadrados ordinarios (MCO). Siguiendo a Arellano (1990), tenemos como resultado el siguiente estimador de efectos fijos de las variables explicativas, llamado también estimador “intra-grupos”:

$$\hat{\beta}_{EF} = (X' \bar{Q} X)^{-1} X' \bar{Q} y = (\tilde{X}' \tilde{X})^{-1} \tilde{X}' \tilde{y} \quad (3)$$

En donde: $\bar{Q} = I_{NT} - C(C'C)^{-1}C'$ tiene dimensión $NT \times NT$ con $C = I_N \otimes \boldsymbol{\iota}$ es $NT \times 1$, $\tilde{X} = \bar{Q}X$ es $NT \times P$ y $\tilde{y} = \bar{Q}y$ es $NT \times 1$.

En el segundo caso, cuando se supone que α_i es independiente de x_{it} es conveniente realizar la estimación bajo inferencia incondicional lo que se denomina modelo de panel de “efectos aleatorios”. En tal caso, α_i no es un valor fijo, sino que es un componente aleatorio que forma parte del término de perturbación. Por lo tanto, este modelo contiene un término de perturbación

compuesto $u_{it} = \alpha_i + \varepsilon_{it}$. La estimación de los modelos de efectos aleatorios se realiza mediante el método de mínimos cuadrados generalizados (MCG), dando como resultado el siguiente estimador de efectos fijos de las variables explicativas o “intra-grupos”:

$$\hat{\beta}_{MCG} = \left(X' \left[I_N \otimes \hat{\Omega}^{-1} \right] X \right)^{-1} X' \left[I_N \otimes \hat{\Omega}^{-1} \right] y \quad (4)$$

En donde Ω es la matriz de covarianzas, cuyos elementos son calculados mediante:

$$\hat{\omega}_{is} = \frac{1}{N} \sum_{i=1}^N \hat{u}_i \hat{u}_{is} \quad (4.1)$$

En la práctica, para determinar si el efecto individual y las regresoras observadas están correlacionas, y por ende realizar la elección del modelo más correcto, es común utilizar el test de especificación de Hausman (1978). Esta prueba de hipótesis estadística permite evaluar si un estimador $\hat{\beta}_e$ que es más eficiente es también consistente comparándolo con otro estimador alternativo $\hat{\beta}_c$ que se sabe que es consistente. Esto se alcanza evaluando si las diferencias entre las estimaciones de ambos modelos son o no sistemáticas. En el caso de que no se presenten diferencias sistemáticas, ambos estimadores serían consistentes y, por lo tanto, sería mejor estimador $\hat{\beta}_e$ que es además más eficiente.

La prueba de Hausman queda definida entonces como:

$$h = \left[\hat{\beta}_{EF} - \hat{\beta}_{MCG} \right]' \left[\text{Var} \left(\hat{\beta}_{EF} \right) - \text{Var} \left[\hat{\beta}_{MCG} \right] \right]^{-1} \left[\hat{\beta}_{EF} - \hat{\beta}_{MCG} \right] \quad (5)$$

Considerando el cumplimiento de los supuestos intrínsecos de cada modelo, el estimador de “efectos fijos” $\hat{\beta}_{EF}$ es consistente, mientras que, el estimador de “efectos aleatorios” calculado por mínimos cuadrados generalizados $\hat{\beta}_{MCG}$ es más eficiente (asintóticamente) pero inconsistente cuando el modelo está mal especificado. Esto es $Var[\hat{\beta}_{MCG}] \leq Var(\hat{\beta}_{EF})$.

IV.2. MODELO DE PSEUDO-PANEL.

En los modelos de panel una de las limitaciones que afectan a su estimación es la disponibilidad y contenido de los datos. Esto es, la presencia de errores de medida debido a la no respuesta o falsa respuesta en encuestas (comunes en las variables de ingresos), los paneles rotativos (muestreo en las encuestas de hogares) donde el panel está un tiempo relativamente corto en la muestra lo que imposibilita su seguimiento y el efecto del “desgaste” o abandono no aleatorio de individuos.

El enfoque de pseudo panel fue inicialmente presentado por el premio nobel Angus Deaton en 1985 con el objetivo de superar estas últimas limitaciones mencionadas mediante la construcción de paneles "sintéticos". Esto se logra reemplazando las observaciones individuales del panel original con medias de subgrupos de la población que se puede identificar su aparición en repetidas encuestas transversales (Meng Y, y otros, 2014).

Para definir los subgrupos se deben seleccionar variables-factores que deben ser o suponerse invariantes en el tiempo (por ejemplo, año de nacimiento, género, etnia). A su vez, garantizar que las cohortes se construyan sobre una población y criterios estables que permitan que sus perfiles no cambien bruscamente en el tiempo.

A partir de la selección de las variables-factores, se agregan aquellos individuos “i” pertenecientes a la cohorte “c” que se observan en la muestra en cada momento “t”.

Así, tenemos las cohortes definidas $c = 1, \dots, C$. donde n_c ³² es el número de observaciones dentro de la cohorte “c”. El tamaño de las cohortes de individuos n_c es importante, y depende directamente de la cantidad de cohortes (C) que se establezcan. Un número mayor de observaciones en cada cohorte garantiza mayor consistencia en la estimación, pero a su vez implica mayor heterogeneidad dentro del subgrupo. Por lo tanto, la elección de n_c mediante incorporación de variables factores³³ genera un “*trade off*” entre homogeneidad y robustez.

El conjunto de individuos (N) en un momento “t” queda definido como $N = C * n_c$. Mientras que, el conjunto de datos en secciones repetidas (S) satisface: $S = C * n_c * T * P = N * T * P$.

En líneas generales, lo que intentamos obtener es la esperanza de la variable de interés para la cohorte “c” en el momento “t”. Esto es $y_{ct}^* = E(y_{it} / i \in c, t)$. Mientras que cada variable predictora será también una esperanza condicional de la cohorte en cada momento, es decir, para cada variable “p” tenemos $z_{ct}^* = E(z_{it} / i \in c, t)$.

Partiendo del modelo de panel estándar presentado en (1), el modelo de pseudo panel satisface la siguiente relación (M. Guillerm, 2017):

³² En este caso se supone que las n_c son iguales.

³³ La incorporación de cada variable factor implica una restricción que debe cumplir la observación para integrar la cohorte. Al incorporar más factores, o factores con categorías más específicas, los subgrupos se vuelven más homogéneos en sus características. Sin embargo, esto limita la posibilidad que las observaciones individuales cumplan con dichos requisitos por lo que se generan subgrupos más pequeños.

$$y_{ct}^* = x_{ct}^* \beta + \alpha_c^* + \varepsilon_{ct}^* \quad (6)$$

$$c = 1, \dots, C. \quad t = 1, \dots, T.$$

Sin embargo, en la práctica, los valores verdaderos tanto de y_{ct}^* como de x_{ct}^* para las cohortes poblacionales son inobservables, por lo que se estiman mediante las cohortes muestrales

$$\text{observadas mediante } \bar{y}_{ct} = \frac{1}{n_c} \sum_{i=1}^{n_c} y_{it} \text{ y } \bar{x}_{ct} = \frac{1}{n_c} \sum_{i=1}^{n_c} x_{it}.$$

En este sentido, el modelo de pseudo panel para medias de muestras de cohortes observadas se expresa como:

$$\bar{y}_{ct} = \bar{x}_{ct} \bar{\beta} + \bar{\alpha}_c + \bar{\varepsilon}_{ct} \quad (7)$$

En donde \bar{y}_{ct} es la media de variable de interés para la cohorte “c” en el momento “t” y $\bar{\alpha}_c$, es la media de los efectos fijos a nivel cohorte para aquellos miembros de la muestra.

La obtención de vector β se realiza a partir del centrado de cada cohorte respecto al valor medio observado $(\bar{y}_{ct} - \bar{y}_c)$ y la posterior aplicación del método de mínimos cuadrados ordinarios (MCO). Siguiendo a Guillerm, (2015) tenemos como resultado el siguiente estimador de efectos fijos o “intra-grupos”:

$$\hat{\beta}_{pp} = \left[\sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c)' (\bar{x}_{ct} - \bar{x}_c) \right]^{-1} \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c)' (\bar{y}_{ct} - \bar{y}_c) \quad (8)$$

El estimador del efecto fijo no observado para la media de la población de cohorte, α_c , deducido a partir de la obtención de $\hat{\beta}_{pp}$ es igual a :

$$\hat{\alpha}_c = \bar{y}_c - \bar{x}_c \hat{\beta}_{pp} \quad (9)$$

El enfoque de pseudo panel permite superar, o al menos atenuar, muchas de las dificultades que se presentan en los modelos de panel. Posibilita el seguimiento de cohortes a lo largo del tiempo en secciones transversales repetidas, generando series de tiempo para las medias de los subgrupos, que se pueden usar como si los datos de panel estuvieran disponibles; atenúa en gran medida el sesgo derivado de los errores de medida (Atman y McKenzie, 2005) y si el tamaño de las cohortes es suficientemente grande ($n_c \rightarrow \infty$) entonces $\bar{\varepsilon}_{ct} = \frac{1}{n_c} \sum_{i=1}^{n_c} \varepsilon_{it} \xrightarrow{P} E(\varepsilon_{it}) = 0$ (Perera, 2006) por lo que asegura parámetros consistentes (Meng Y, y otros, 2014).

Con los modelos de pseudo panel se logran buenos resultados (Deaton, 1985). Sin embargo, presentan algunas limitaciones, como que la submuestra de individuos para estimar los verdaderos valores de y_{ct}^* como la de x_{ct}^* puede no ser representativa (Guillerm, 2017). Además, es necesario que los tamaños de muestra de la cohorte sean lo suficientemente grandes para que el promedio de los efectos fijos de la muestra $\bar{\alpha}_{ct}$ sean una buena aproximación del efecto fijo medio poblacional no observado de la cohorte α_{ct}^* (Deaton, 1985). Dado que los individuos observados en cada momento no son los mismos, el promedio de efectos fijos $\bar{\alpha}_{ct}$ puede variar cuándo teóricamente debe ser constante (Guillerm, 2017). Aunque el modelo de pseudo panel puede reducir la endogeneidad, la misma sigue siendo un problema (Meng Y, y otros, 2014).

IV.3. METODO DE CLASIFICACIÓN.

El análisis de conglomerados o clustering típico es una técnica que tiene por objetivo la agrupación o clasificación de un conjunto de datos en grupos (clústers) de objetos similares. En consecuencia, los clústers formados son agrupamientos de datos cuyo criterio exige que los

elementos dentro de cada grupo sean más homogéneos, en término de las características analizadas, en relación a los objetos de otros grupos (Garre, Cuadrado, Sicilia, Rodríguez & Rejas, 2007). Para medir la similitud entre estos objetos, en función de las variables involucradas, se utilizan métricas matemáticas que no contemplan el aspecto temporal.

Por su parte, las técnicas de clúster-longitudinal combinan, en una sola representación, similitudes de contenido y adyacencia temporal. Esto implica utilizar algoritmos que tengan en cuenta la vecindad temporal de los objetos para extraer conocimiento útil y poder realizar un análisis lo más completo posible de la dinámica y estructura de la problemática objeto de estudio, como así también, la relación entre sus múltiples factores y determinantes. C. Genolini y B Falissard (2016) y C. Genolini, B. Falissard y J-B. Pingault (2017) desarrollaron el algoritmo no paramétrico k-means que permite trabajar con trayectorias simples o conjuntas³⁴.

Sea S un conjunto de “ C ” sujetos (cohortes de individuos o pseudo paneles) sobre los que se miden “ P ” atributos o variables en los “ T ” momentos de tiempo. Esto es:

$$S = C * T * P \tag{10}$$

En donde cada una de las cohortes de individuos es representado por “ c ” siendo $c=1, \dots, C$, y

“ C ” el número total de cohortes $C = \frac{N}{n_c}$ para cohortes del mismo tamaño. El subíndice $t=1, \dots, T$,

indica cada uno de los momentos del tiempo en donde se realizan las mediciones de las P variables de interés y_1, \dots, y_p sobre las observaciones. Mientras que, subíndice “ p ” de y_p indica el atributo que se está midiendo $p=1, \dots, P$.

³⁴ El algoritmo esta implementado en el lenguaje R con el paquete “k-means for Longitudinal data” desarrollado por los autores, el cual se utilizó en el presenta trabajo, con las funciones “klm” para trayectorias simples y “kml3” para trayectorias conjuntas.

La secuencia conformada por las mediciones de “p” sobre una cohorte “c” en los “T” momentos temporales $y_{c,p} = (y_{c1p}, y_{c2p}, \dots, y_{cTp})$ se denomina “trayectoria simple” de la variable “p” en la cohorte “c”.

En consecuencia, la sucesión de “p” en los “C” pseudo paneles en los “T” periodos de tiempo $Y_{..p}$ constituye la “trayectoria simple” de la característica “p” para el conjunto de observaciones

$$C = \frac{N}{n_c}.$$

Por su parte, la serie de las “P” variable trayectoria simple $Y_{..1}, Y_{..2}, \dots, Y_{..P}$ calculadas en “T” tiempos sobre los “C” subgrupos se denomina como “trayectoria conjunta” para el conjunto de observaciones “C”.

Por lo tanto, podemos definir a cada sujeto “c” como la matriz $Y_{c..}$ con dimensión P x T de su “trayectoria conjunta” $Y_{c,1}, Y_{c,2}, \dots, Y_{c,P}$.

$$Y_{c..} = \begin{pmatrix} y_{c11} & y_{c21} & \cdots & y_{cT1} \\ y_{c12} & y_{c22} & \cdots & y_{cT2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{c1P} & y_{c2P} & \cdots & y_{cTP} \end{pmatrix} \quad (11)$$

$c=1, \dots, C.$

$t=1, \dots, T.$

$p=1, \dots, P.$

Las filas de la matriz (11) $Y_{c,p} = (y_{c1p} \quad y_{c2p} \quad \dots \quad y_{ctp})$ indican cada una de las P “trayectoria simple” en la cohorte “c”. Es decir, cada una de las P filas refleja la evolución temporal de los atributos o características analizadas para cada sujeto.

Las columnas de la matriz (14) $Y_{ct.} = \begin{pmatrix} y_{ct1} \\ y_{ct2} \\ \vdots \\ y_{ctp} \end{pmatrix}$ indican el “Estado de la cohorte c” en cada uno de

los t momentos temporales. Es decir, cada una de las T columnas indican la situación o el estado de los atributos de cada sujeto en un momento determinado del tiempo, tal como en los análisis transversales.

El conjunto de datos en secciones repetidas (S) se compone de C matrices $Y_{c..}$.

Esto es: $S = C * Y_{c..}$

Nuestro objetivo es dividir al conjunto S en M sub-grupos de cohortes de individuos. La condición que debe cumplir el agrupamiento es que, en término de las características analizadas temporalmente, cada uno de los “m” sub-grupos sean conformado de manera tal que las cohortes que lo integren sean más homogéneas respecto a las cohortes que integran a los otros sub-grupos.

En donde $m=1, \dots, M$.

Para medir la similitud o distancia “d” entre dos cohortes de individuos $Y_{1..}$ e $Y_{2..}$ se define una métrica o función distancia *Dist* y se utiliza la norma euclidiana $\|\cdot\|$.

Siguiendo a Genolini & otros (2015), el enfoque “k-means for Longitudinal data” (kml) plantea dos métodos para medir esta distancia “ \mathbf{d} ”. El primer método calcula la distancia “ d^t ” entre dos observaciones $Y_{1..}$ e $Y_{2..}$ teniendo en consideración el “Estado de las cohortes” en cada momento “ t ” mientras que el segundo método calcula la distancia “ d^p ” considerando las “trayectoria simple” para cada característica “ p ”.

Así pues, en el primer método, para cada tiempo “ t ” fijo definimos la distancia “ d^t ” entre $Y_{1..}$ e $Y_{2..}$ como $d_t(Y_{1t}, Y_{2t}) = Dist(Y_{1t}, Y_{2t})$. Esto es, la distancia entre los “estados de las cohortes” en el momento “ t ”, o sea la distancia entre la columna “ t ” de la matriz $Y_{1..}$ y la columna “ t ” de la matriz $Y_{2..}$.

El resultado es un vector de T distancias entre las dos cohortes de individuos:

$$(d_1(Y_{11}, Y_{21}), d_2(Y_{12}, Y_{22}), \dots, d_T(Y_{1T}, Y_{2T}))$$

Combinando las T distancias mediante la función de la norma $\|\cdot\|$ se obtiene la distancia entre $Y_{1..}$ e $Y_{2..}$ bajo este primer método

$$d^t(Y_{1..}, Y_{2..}) = \|(d_1(Y_{11}, Y_{21}), d_2(Y_{12}, Y_{22}), \dots, d_T(Y_{1T}, Y_{2T}))\| \quad (12)$$

En el segundo método, para cada variable “ p ” fija definimos la distancia “ d^p ” entre $Y_{1..}$ e $Y_{2..}$ como $d_p(Y_{1.p}, Y_{2.p}) = Dist(Y_{1.p}, Y_{2.p})$, es decir, la distancia entre la fila “ p ” de la matriz $Y_{1..}$ y la

fila “p” de la matriz $Y_{2..}$. Esto es, la distancia entre la “trayectoria simple” o la evolución temporal de la variable “p” entre los sujetos.

El resultado es un vector de P distancias entre las dos cohortes de individuos:

$$(d_{.1}(Y_{1.1}, Y_{2.1}), d_{.2}(Y_{1.2}, Y_{2.2}), \dots, d_{.p}(Y_{1.p}, Y_{2.p}))$$

Combinando las P distancias mediante la función de la norma $\|\cdot\|$ se obtiene la distancia entre $Y_{1.}$ mediante el segundo método:

$$d^p(Y_{1.}, Y_{2.}) = \left\| (d_{.1}(Y_{1.1}, Y_{2.1}), d_{.2}(Y_{1.2}, Y_{2.2}), \dots, d_{.p}(Y_{1.p}, Y_{2.p})) \right\| \quad (13)$$

Si la medida de distancia corresponde a la métrica de Minkowski se obtienen distancias equivalentes para ambos métodos, $d^l(Y_{1.}, Y_{2.}) = d^p(Y_{1.}, Y_{2.})$ siendo para el primer método

$$\sqrt[a]{\sum_{t=1}^T |Y_{1tp} - Y_{2tp}|^a} \quad \text{y para el segundo método} \quad \sqrt[a]{\sum_{p=1}^P |Y_{1tp} - Y_{2tp}|^a} .$$

En este sentido, corresponde a nueva implementación de k-medias planteado específicamente para trabajar con datos de esta naturaleza, admite buenos resultados en trayectorias no polinómicas y no exige supuestos forzados en la distribución (Genolini, C., & Falissard, B., 2010).

En términos computacionales, el k-medias es un algoritmo de optimización de Escalada Simple (o “hill climbing”) que pertenece al grupo de Esperanza-Maximización. Inicialmente, cada observación es asignada a un grupo. Luego, el agrupamiento óptimo se alcanza mediante la alternancia de dos fases: Una primera fase donde se calculan los centroides de cada grupo

(semillas) y una segunda fase donde la maximización consiste en la asignación de cada observación al grupo más cercano. Esta alternancia entre etapas se repite hasta que no se produzcan más cambios en los grupos (G. Celeux and G. Govaert, 1992).

Partiendo de una configuración inicial, el algoritmo converge hacia un máximo, pero no hay garantías respecto de que este sea global o local. Una solución parcial, que ofrece el paquete utilizado es que ejecuta automáticamente el algoritmo descrito, variando las condiciones de inicio para un número diferente de agrupamientos buscado. Esto permite realizar múltiples sorteos (“*rerolling*”) para cada uno de estos agrupamientos elegidos que conducirán hacia el máxima global (Genolini, C., Alcoque X. & Otros, 2015).

La elección óptima de la partición resultante se obtuvo evaluando los criterios de selección más frecuentes considerando diferentes medidas de disimilitud en los sucesivos sorteos.

Sea c_m la cantidad de cohortes en el grupo “m”, entonces la partición de los C elementos se puede expresar como $C = c_m \times M$, bajo el supuesto que son iguales.

Se define a la matriz de suma de cuadrados entre grupos (B) como:

$$B = \sum_{m=1}^M c_m (\bar{y}_m - \bar{y})(\bar{y}_m - \bar{y})' \quad (14)$$

donde \bar{y} es la trayectoria media de la variable “y” en el conjunto \bar{y}_m denota la trayectoria media de la variable “y” del cluster “m”.

Mientras que, la matriz de suma de cuadrados dentro de los grupos (W) es:

$$W = \sum_{m=1}^M \sum_{k=1}^{n_m} (y_{mk} - \bar{y}_{mk})(y_{mk} - \bar{y}_{mk})' \quad (15)$$

Donde y_{mc} representa trayectoria de “y” para la cohorte “c” en el grupo “m”.

Los criterios para medir homogeneidad, plantean en general la minimización de la matriz W. Esto es, al minimizar la variabilidad de las trayectorias se obtienen grupos más homogéneos. Mediante el mismo razonamiento, la maximización de matriz B nos permitiría obtener grupos más diferenciados entre sí.

Para determinar el número óptimo de grupos se pueden seguir varios criterios, de los que se seleccionan el criterio de T. Caliński y J. Harabasz (1974).

El criterio de Calinski y Harabasz para M grupos maximiza la siguiente relación:

$$CH(M) = \frac{\text{traza}(B) \ c - M}{\text{traza}(W) \ M - 1} \quad (16)$$

PARTE V. RESULTADOS.

El quinto capítulo presenta la aplicación de las técnicas descritas en el apartado de metodología. En tal sentido, se detallan los conceptos y elementos referidos a su procesamiento, y la posterior lectura de los resultados. En la primera sección (V.1), se delimitan las características específicas de la estructura longitudinal del conjunto de datos. En la segunda sección (V.2), se hace referencia a los software y paquetes estadísticos utilizados, tanto en la modelación de pseudo panel, como en la clasificación longitudinal. En la sección V.4 y V.5, se exponen e interpretan los resultados del modelo de pseudo panel y clustering longitudinal, respectivamente. Finalmente, en la última sección (V.6), se describen las características sociodemográficas y económicas de los tres grupos de cohortes resultantes, en donde se evalúa la interseccionalidad según género y generación en cada uno de éstos. Los resultados indican la presencia brecha de género y efecto cohorte; como así también, existencia de segmentación sociodemográfica en la estructura del mercado laboral argentino.

V.1. ESTRUCTURA DEL CONJUNTO DE DATOS.

El conjunto de datos a partir del cual se procesó la información y se realizaron los posteriores análisis corresponde a la “Base de datos de Trabajo”, y construida a partir de las Bases Usuarias de la Encuesta Permanente de Hogares.

La base de datos de trabajo, descrita en el apartado II.3.2, contiene como unidades de observación a las “C” cohortes de trabajadores asalariados pertenecientes a los grandes centros urbanos de Argentina en el período 1990-2003. Cada cohorte fue definida por el conjunto de trabajadores en relación de dependencia que comparten simultáneamente el año calendario de nacimiento y el género declarado. Los atributos de la base de datos usuaria original se consideran como la media estimada para cada cohorte.

En el presente trabajo, el análisis del conjunto de datos es limitado a las cohortes de trabajadores asalariados varones y mujeres que durante los años en estudio presentaron edad comprendidas entre los 18 y 60 años. El límite inferior corresponde a la edad mínima que admite, en términos generales, la legislación laboral Argentina³⁵ para trabajar en relación de dependencia y coincide con la edad modal de finalización del nivel de instrucción educativa media en el país. Mientras que, el límite superior referencia a la edad laboral máxima, en términos generales, de trabajadoras asalariadas y trabajadores en tareas especiales o de riesgo.

En consecuencia, la base de datos utilizada para el procesamiento y análisis se compone de 25 variables y 1204 observaciones correspondientes a 112 pseudo-paneles.

³⁵ La ley N° 20.744 de Régimen de Contrato de Trabajo establece que a partir de los dieciocho años las personas pueden celebrar contrato de trabajo. Las personas mayores de dieciséis años no emancipadas pueden celebrar contrato de trabajo solamente con la autorización de sus padres, responsables o tutores siempre que no corresponda a tareas de riesgo. Mientras que, las personas mayores de catorce años solo pueden ser ocupados en empresas cuyo titular sea su padre, madre o tutor en jornada reducida siempre que no realice a tareas de riesgo.

En el conjunto de datos de panel, ciertas cohortes carecen de información en algunos de los catorce años que componen el estudio. Esto se debe a que determinadas cohortes ingresan al mercado laboral en años posteriores al año inicial (1990) y otras superan la edad máxima propuesta de 60 años en el trayecto del periodo saliendo del grupo de análisis. En este caso, tal como indica la salida de Stata (Tabla D.2), se denomina al conjunto de datos como panel no balanceado.

La estimación en modelos de panel no balanceados con efectos fijos, como el que consideramos a continuación, no presenta importantes diferencias en relación a los modelos con paneles balanceados.

Si el panel fuera balanceado como se consideró en el capítulo IV, se tienen “C” sujetos (cohortes de individuos o pseudo paneles) en “T” diferentes momentos de tiempo, y donde cada de las $c = 1, \dots, C$. cohortes participa en los $t = 1, \dots, T$. momentos de tiempo. Entonces, la totalidad de observaciones en el conjunto de datos de panel se puede expresar mediante el cardinal³⁶

$$N^c = C * T \quad (17)$$

En el caso de conjuntos con paneles no balanceados el número total de observaciones es entonces

$$T_1 + T_2 + \dots + T_{N^c}.$$

Donde T_c para $c=1, 2, 3 \dots N$ es el número de períodos de tiempo en que la cohorte “c”. Así por ejemplo, si la cohorte “c” en todos los años que van de 1990 a 2003 presentó edades comprendidas entre los 18 y 60 años, entonces el cardinal T_c asumirá el valor 14.

³⁶ Si suponemos que en el conjunto de datos existe un solo atributo o variable de medición. Es decir que, en términos de la ecuación 10 del capítulo IV, P asumen el valor 1, entonces, la totalidad de observaciones del conjunto de datos de panel coincide con el tamaño del conjunto de datos $S = C * T * P$.

La pérdida en grados de libertad y el procedimiento de la regresión de variables binarias del efecto fijo es análoga en paneles balanceados o no balanceados. En tal sentido, aquellas cohortes que participan en solo un momento del tiempo no realizan aporte al análisis de efectos fijos dado que la diferencia entre periodos (ΔT_c) no se puede aplicar. Esto último, no afecta el estudio realizado dado que excluye del análisis a aquellas cohortes que no cumplen con el principio de longitudinalidad.

V.2. SOFTWARE Y PAQUETES.

El procesamiento de datos constó de tres etapas: En una primera instancia se identificaron las cohortes, bajo los criterios establecidos, en las catorce Bases Usuarías de la Encuestas Permanentes de Hogares Puntual correspondientes a la primera onda de los años comprendidos entre 1990 y 2003 para cada uno de los siete aglomerados urbanos en estudio. La posterior unificación de las mismas se realizó previo trabajo de homogeneización de las variables relevantes al análisis y la elaboración de los indicadores de interés. Luego de la unificación de las bases de datos en un conjunto que contemplaba la dimensión temporal se seleccionaron solo aquellas observaciones referentes a trabajadores cuya categoría ocupacional era la asalariada y se calcularon las medias de los atributos sobre los agregados correspondientes a cada una de las cohortes o pseudo paneles.

En una segunda instancia, se retuvo del conjunto de datos de pseudo-panel solo aquellas observaciones que presentaron en los años de referencia edades comprendidas entre 18 y 60 años. A partir de esta selección se evaluaron distintos modelos de panel para estimar los niveles de informalidad en cada cohorte y el ajuste predicho.

Tanto la primera etapa que contenía la depuración, medición de atributos y construcción de base de trabajo como la segunda etapa de modelación de datos y análisis de posestimación, fueron realizadas por medio del paquete de software estadístico Stata Portable versión 14.

En la última etapa, se aplicó la técnica de clasificación no supervisada longitudinal “K_mL” en los valores medios de los subgrupos de trabajadores, y también, sobre los valores predichos para cada una de estas cohortes resultantes del modelo de pseudo-panel elegido en la instancia previa. La elección de la clasificación más adecuada se efectuó previo ensayo de distintas métricas y el contraste de distintos criterios de selección. Esto último, se ejecutó mediante entorno de desarrollo RStudio Server versión 1.4.17 integrado para el lenguaje libre de programación R versión 4.1.0. El algoritmo de clasificación corresponde a los paquetes K-Means para datos longitudinales (“kml”) versión 2.4.1³⁷ y K-Means para datos longitudinales conjuntos (“kml3d”) versión 2.4.2³⁸. Ambos paquetes se encuentran contenidos en el repositorio CRAN³⁹ de R Project.

V.3. MODELO DE PSEUDO PANEL: RESULTADOS.

Una gráfica de las trayectorias de la informalidad laboral en cada una de las cohortes generacionales de trabajadores asalariados por sexo representa una primera aproximación al enfoque longitudinal de los datos (Las Figuras D.1 del Anexo D). Las cohortes que en el año base poseían edades comprendidas en el rango central de las consideradas económicamente activas, presentaron menor volatilidad en sus niveles medios de informalidad laboral y no exhiben una tendencia temporal marcada como sucede con las restantes. En términos generales, en las edades

³⁷ El paquete “K-Means for Joint Longitudinal Data” (2016) es de acceso libre y tiene por autores a C. Genolini y B. Falissard.

³⁸ El paquete “K-Means for Joint Longitudinal Data” (2017) es de acceso libre y tiene por autores a C. Genolini, B. Falissard y J.-B. Pingault.

³⁹ CRAN es la sigla de Comprehensive R Archive Network para el lenguaje de programación R. Los paquetes contenidos en este repositorio son libres y se encuentran en <https://cran.r-project.org/>.

de ingreso al mercado de trabajo, como así también, en edades de actividad avanzadas es notoria la presencia de tendencia en la serie. Sin embargo, en el primer caso, se observa una menor pendiente en cohortes que ingresaron al mercado laboral en años posteriores al año de inicio del estudio. Esto último, se traduce en una transición de incorporación laboral que requirió una mayor cantidad de años para alcanzar los niveles de formalidad medios del período.

Tabla V.1
Prueba de Hausman.

Variable	Coeficientes		Diferencia (b-B)	S.E. sqrt(diag(V_b-V_B))
	Fijo (b)	Aleatorio (B)		
Secundario				
Industria	-0.15	-0.14	-0.01	0.01
Construccion	0.06	-0.02	0.08	0.01
Intensidad	0.15	0.07	0.08	0.04
antig	0.16	0.24	-0.08	0.02
Tam_Estab	-0.32	-0.37	0.05	0.02
t	0.32	0.39	-0.07	0.01
REdad	0.03	0.01	0.01	0.00
Etapa	-0.17	-0.01	-0.16	0.05
2	-0.05	-0.05	0.00	
3	-0.09	-0.09	0.00	

Prueba

b = Consistente bajo Ho y Ha.

B = Inconsistente bajo Ha, eficiente bajo Ho

Ho : Diferencia en coeficientes no sistemática.

H1 : Diferencia en coeficientes sistemática.

Resultado

$$\chi^2_{(10)} = (b-B)'[(V_b-V_B)^{-1}](b-B) = 65.69$$

$$\text{Prob} > \chi^2 = 0.0000$$

A partir de los datos del pseudo-panel se propone un modelo para determinar los factores que explican el nivel de informalidad laboral en las cohortes de trabajadores asalariados de los principales centros urbanos de Argentina. El conjunto de datos consta de 1204 observaciones pertenecientes a 112 cohortes de trabajadores mujeres (56) y varones (56) cuya participación promedio es próxima a 11 de los 14 años en estudio.

Para considerar en el modelo autocorrelación propia de los modelos de panel se incorpora al modelo el efecto cohorte. Un primer paso consiste en determinar si el modelo corresponde a un modelo de efectos fijos o efectos aleatorios, para lo cual se definió la prueba de especificación de Hausman, tal como se indica en el apartado IV.1. Los resultados obtenidos indican el rechazo de la hipótesis nula del ensayo. En consecuencia, las diferencias entre los coeficientes de ambos modelos son sistemáticas y el estimador de efectos aleatorios no es consistente. Por lo tanto, lo adecuado es la estimación de efectos fijos mediante mínimos cuadrados ordinarios.

La determinación de modelar mediante efectos fijos o efectos aleatorios se definió, tal como se indica en el apartado IV.1, mediante la prueba de especificación de Hausman. Los resultados obtenidos indican el rechazo de la hipótesis nula del ensayo. En consecuencia, las diferencias entre los coeficientes de ambos modelos son sistemáticas y el estimador de efectos aleatorios no es consistente. Por lo tanto, lo adecuado es la estimación de efectos fijos mediante mínimos cuadrados ordinarios.

Tabla V.2

*Modelo de pseudo-panel con efectos fijos (dentro de grupos).
Principales pruebas del modelo.*

Número de observaciones = 1204

Número de grupos = 112

Variable de Grupo : id

R- Cuadrado		Observaciones por grupo	
Dentro	= 0.5542	Mínimo	= 2
Entre	= 0.7843	Promedio	= 10.8
General	= 0.5397	Máximo	= 14

corr (U_i, X_b) = -0.8572

F(10,1082) = 114.54

Prob > F = 0.0000

Prueba F todos los U_i = 0 :

F(111, 1082) = 1.89

Prob > F = 0.0000

Modelo de pseudo-panel. Coeficientes y estadísticos asociados.

Informal	Coef.	E. E.	t	P > t	[I. C. 90%]	
					LI	LS
Secundario	-0.15	0.03	-5.56	0.0000	-0.19	-0.10
Industria	0.06	0.03	1.72	0.0850	0.00	0.11
Construccion	0.15	0.07	2.08	0.0380	0.03	0.27
Intensidad	0.16	0.03	5.05	0.0000	0.11	0.21
antig	-0.32	0.03	-11.59	0.0000	-0.37	-0.28
Tam_Estab	0.32	0.03	12.16	0.0000	0.28	0.37
t	0.03	0.00	6.31	0.0000	0.02	0.04
REdad	-0.17	0.05	-2.95	0.0030	-0.25	-0.08
Etapa						
2	-0.05	0.01	-6.62	0.0000	-0.07	-0.04
3	-0.09	0.01	-6.68	0.0000	-0.12	-0.07
Constante	1.45	0.26	5.01	0.0000	0.99	1.91
σ_u	0.1994					
σ_e	0.0629					
ρ (rho)	0.9094	(fraction of variance due to u _i)				

Fuente: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual.

La Tabla V.2 muestra los resultados obtenidos del modelo de pseudo-panel con efectos fijos (o de covarianza), en su parte superior muestra un resumen de los datos. El coeficiente de determinación (R^2) se descompone según variabilidad intra-grupo, entre-grupos, y general. El primero de estos, indica que a nivel intra-cohorte (dentro), el modelo seleccionado explica al 55,4 por ciento de las variaciones en la informalidad laboral. La bondad de ajuste del modelo para explicar las variaciones de la variable dependiente entre las cohortes (entre) alcanza el 77,4 por ciento. A su vez, considerando el conjunto de observaciones de las cohortes en general, la proporción de la varianza total de la variable explicada por la regresión es próximo al 54 por ciento alcanzando un importante poder explicativo.

El tercer recuadro de la parte superior de la Tabla V.2 “*Principales pruebas del modelo*” indica que los errores u_i se encuentran fuertemente correlacionados y de forma negativa (-0.85) con los regresoras del modelo de efectos fijos. Esta potencial correlación entre la variable no observada (u_i) con las regresoras observadas, refuerza la elección de un modelo efectos fijos frente a un modelo de efectos aleatorios. Por su parte, la prueba F general que establece que todos los estimadores de este modelo son iguales a cero es rechazada al nivel de significación del 5 por ciento.

La prueba de hipótesis presentada en el último recuadro plantea el supuesto de que todas las variables dicotómicas que indican los efectos fijos son iguales a cero. El estadístico F observado resultante se ubica en la zona de rechazo de la distribución concluyendo el rechazo de la hipótesis nula definida. En conclusión, los primeros resultados presentados indican que el modelo general es adecuado en su determinación, análisis de varianza y existencia de efecto-cohorte.

Para las variables regresoras en la parte inferior de la Tabla V.2 se muestra los coeficientes, su error estándar, el intervalo de confianza de los estimadores obtenidos el valor del estadístico de la prueba t de significancia con su respectivo valor de probabilidad (p-value). En la parte inferior de la tabla se presenta la descomposición de la varianza. entre variabilidad inter e intraclase, (u_i desvíos estándar de los residuos dentro las cohortes ϵ_i los desvíos estándar de los residuos generales y ρ el coeficiente de correlación interclase).

Las variables “Comercio”, “Primario” y “IngHora”, analizadas en la Parte III de este trabajo, no fueron incluidas entre las regresoras,” Comercio” mide la proporción de trabajadores en las cohortes que se encuentran empleados en actividades económicas relacionadas al comercio mayorista o minorista, resultó no significativa al nivel del 10 por ciento. “Primario” que indica el promedio de trabajadores con nivel educativo primario completo en la cohorte, fue omitida por presentar colinealidad con la variable que indicaba la media de trabajadores en la cohorte que finalizaron la educación media (secundaria) o un nivel de educación equivalente. Por último, “IngHora” es el ingreso medio percibido por la cohorte en su actividad laboral principal, y fue excluido debido a la potencial causalidad inversa o simultanea entre esta variable y la variable dependiente.

Las variables regresoras que resultaron significativas fueron: industria (a un nivel de significación del 10 por ciento), construcción (a un nivel de significación del 5 por ciento) y; la constante, Secundario, Intensidad, Antigüedad, Tamaño de Establecimiento, tiempo, Edad (raíz), Etapa temporal a un nivel de significación del 1 por ciento.

Los coeficientes estimados indican que las variables más relevantes para explicar las variaciones en los niveles de informalidad laboral en las cohortes de trabajadores asalariados fueron la

antigüedad y el tamaño del establecimiento empleador. En tal sentido, el modelo indica que, ante un aumento unitario en la proporción de empleados con antigüedad mayor a un año en la cohorte, su tasa de informalidad disminuye en promedio en 0.32, mientras que, ante un aumento unitario en la proporción de trabajadores en microempresas en la cohorte, se estima un aumento en la informalidad en la misma magnitud anterior (0.32).

Otras causales de incrementos en el empleo informal de las cohortes de trabajadores, son: el aumento en la proporción de subempleo en la cohorte (0.16), el aumento en la participación de trabajadores con empleos en actividades de industria manufacturera (0.06) y construcción (0.15) en la cohorte y, en términos generales, al aumento en la cantidad de años respecto del año de inicio del estudio (0.03). Por el contrario, los niveles de informalidad disminuyen ante aumentos en la proporción de trabajadores con secundario completo en la cohorte, y en términos generales, el aumento etario (-0.17).

Por su parte, la variable “Etapa” indica el valor diferencial medio de la Etapa 2 (1995-1998) y la Etapa 3 (1999-2003) en relación a la Etapa 1 representado a través de la constante en las etapas.

Las variaciones de los residuos del modelo se dividen entre los desvíos estándar de los residuos del efecto individuo específico invariante en el tiempo (σ_u) y los desvíos estándar de los residuos generales (σ_ϵ). En tal sentido, los resultados en ambos casos sugieren que la parte de la varianza de la variable dependiente que se debe a la heterogeneidad no observada entre cohortes es alta. El 90 por ciento de la variabilidad total se debe a la variabilidad entre cohortes.

La varianza total (σ^2) resultante de la suma de las varianzas de los residuos mencionados

$\sigma^2 = \sigma_u^2 + \sigma_\epsilon^2$ asume el valor 0.044, mientras que, el coeficiente de correlación intraclase –ICC–

$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2}$ es de 0.9094. El ICC indica que más del 90 por ciento de la variabilidad total se

debe a la varianza entre grupos (σ_u). En este sentido, revela que el componente de la varianza a nivel panel es muy importante y, en consecuencia, el estimador de panel obtenido es muy diferente del estimador obtenido para datos agrupados (“pooled”) expresando concordancia en la consistencia del modelo elegido.

Además de los criterios de información y los coeficientes de determinación analizados previamente, la comparación entre distintos modelos, fue definida considerando pruebas de razón de verosimilitud entre dos modelos anidados, de esta manera se determinaron las variables que se excluyeron del modelo (Tabla V.3).

Tabla V.3
Prueba de Razón de Verosimilitud.

Variable excluida	$\chi^2_{(.)}$	gl	Prob > χ^2
Secundario	33.9	1	0.000
Industria	3.3	1	0.049
Construccion	3.8	1	0.042
Intensidad	28.1	1	0.000
antig	140.9	1	0.000
Tam_Estab	154.3	1	0.000
t	43.5	1	0.000
Redad	9.6	1	0.002
i.Etapa	52.1	2	0.000

Modelo 1: Modelo completo.

Modelo 0: Modelo con variable excluida.

Supuesto: Modelo 0 anidado en Modelo 1.

Las Figuras D.2 a D.5 del Anexo D.1 brindan información para realizar el diagnóstico de los supuestos que se evalúan en los modelos lineales sobre los residuos de la estimación.

El histograma (Figura D.2) muestra la aproximación de los errores a la densidad normal presentando un ajuste adecuado, con excepción en la zona central de la distribución, en donde los valores más frecuentes se ubican a la izquierda o derecha del valor promedio. Por último, en el gráfico Q-Q (Figura D.3) los residuos estandarizados se ajustan a la recta de cuantiles teóricos, observándose algunos atípicos en los extremos de la distribución.

Los diagramas de dispersión que relacionan los residuos con el orden de la cohorte (Figura D.4) y los residuos estandarizados con los valores predichos del modelo (Figura D.5) no parecen presentar fuertes patrones que indiquen violación al supuesto de independencia o heterocedasticidad.

V.4. CLUSTERING LONGITUDINAL.

La técnica de clustering longitudinal (“KmL”⁴⁰) seleccionada para este estudio posibilita estudiar la evolución conjunta de las trayectorias de variables medidas en individuos a lo largo del tiempo, es decir, nos permite examinar y clasificar cohortes de trabajadores según la dinámica simultánea de “variables-trayectoria” que constituyen características de interés para determinar segmentos en el mercado de trabajo argentino durante los años que comprende la investigación.

En una segunda etapa, entonces, se clasificó temporalmente a las cohortes en función de dos atributos que determinan la calidad del empleo: La informalidad laboral, de naturaleza cualitativa que materializa las condiciones de trabajo decentes; y el ingreso monetario, de naturaleza cuantitativa que materializa el poder adquisitivo para alcanzar medios de sustento dignos.

⁴⁰ La técnica fue desarrollada por los autores Genolini, C., & Falissard, B. (2010). KmL: k-means for longitudinal data. *Computational Statistics*, 25(2), 317-328.

Los niveles promedios de informalidad, para cada cohorte de trabajadores asalariados, fueron estimados mediante el modelo de efectos fijos (o covarianza) . La modelación realizada permitió “suavizar” la volatilidad de la trayectoria de la variable durante el periodo, que había sido previamente obtenida a partir de valores promedio en cada subgrupo, enfatizando en su tendencia en cada etapa. Asimismo, se incorporó al análisis de clasificación la variable-trayectoria que indica la ubicación de la cohorte en el rango de ingresos medios por hora de cada año, es decir, el nivel de ingreso promedio por hora de la ocupación principal de la cohorte estandarizado en función de los ingresos medios de ese año en el rango cero a uno.

De esta manera en cada momento de tiempo “t” y para cada cohorte “c” fueron inicialmente definidas variables medias $\{\bar{y}_{ct1}, \bar{y}_{ct2}\}$ para la posterior obtención de las variables trayectorias

$\{y_{c...1}, y_{c...2}\}$. Los valores $\bar{x}_{ct} = \frac{1}{n_c} \sum_{i=1}^{n_c} x_{it}$ corresponde a las medias de las variables explicativas en

cada cohorte “c” y momento “t”. Estas variables son: “nivel educativo secundario completo” (x_1), “participación en actividades industriales” (x_2), “participación en actividades de la construcción (x_3), “horas trabajadas semanalmente” (x_4), “antigüedad en el puesto” (x_5), “tamaño de la empresa” (x_6), “tiempo en años” (x_7), y “edad” (x_8)

Estas últimas son resultado de la medición de predichos del modelo de pseudo panel en el primer caso ($y_{c...1}$) y del proceso de estandarización en el segundo caso ($y_{c...2}$).

La trayectoria $y_{c...1}$ se obtuvo a partir de los valores predichos del modelo () $\bar{y}_{ct1} = \frac{1}{n_c} \sum_{i=1}^{n_c} y_{it}$

indica la “tasa de informalidad media” de cada cohorte “c” en el momento “t”

La trayectoria $y_{c,..,2}$ se obtuvo a partir \hat{y}_{ct2} que se estimó en función de su valor medio estandarizado \bar{y}_{ct2} , esto es, $\hat{y}_{ct2} = \bar{\beta} + \bar{\alpha}_c + \bar{\varepsilon}_{ct}$. La variable \bar{y}_{ct2} indica la media del ingreso percibido por hora en la ocupación principal en cada cohorte “c” estandarizado en el rango cero a uno. La estandarización se realizó para evitar problema de escalas entre las variables involucradas. Es decir, para cada cohorte “c” en el momento “t” satisface $\bar{y}_{ct2} = \frac{1}{n_c} \sum_{i=1}^{n_c} y_{it2}^*$. En donde y_{it2} es el ingreso promedio por hora del trabajador “i” en el momento “t” e $y_{it2}^* = \frac{(y_{it2} - \text{mín}(y_{it2}))}{(\text{Máx}(y_{it2}) - \text{mín}(y_{it2}))}$ es el ingreso promedio por hora estandarizado del trabajador “i” en el momento “t”.

Una vez estimadas las trayectorias de los niveles medios de informalidad laboral $y_{c,..,1}$ y el ingreso medio por hora estandarizado $y_{c,..,2}$, cada una de las 112 cohortes de trabajadores es expresada como la siguiente matriz de sus trayectorias conjuntas:

$$Y_{c..} = \begin{pmatrix} y_{c,1,1} & y_{c,2,1} & \cdots & y_{c,13,1} \\ y_{c,1,2} & y_{c,2,2} & \cdots & y_{c,13,2} \end{pmatrix}$$

La medida de disimilitud utilizada para medir el grado de semejanza transversal $d^t(Y_{1..}, Y_{2..})$ y temporal $d^p(Y_{1..}, Y_{2..})$ entre cohortes fue la métrica Euclideana. Se analizaron distintos agrupamientos para k grupos (k =2,3,4,5,6) con 20 corridas (“rerolling”) del algoritmo k-medias en cada agrupamiento, se seleccionó el mejor según el criterio de optimización de Calinski y Harabasz (CH) para los cuales se obtuvo como valor del índice el promedio de los valores resultantes en cada re-lanzamiento del algoritmo en cada número de agrupamiento.

La elección de la métrica euclidiana como función de distancia fue en virtud de que arrojó mejores niveles de optimización en el cociente entre variabilidad entre grupos y variabilidad dentro de grupos. Los resultados expuestos en la Tabla V.4 confirman esto último. Las funciones distancia utilizadas fueron Euclidea, Máxima, Minkowski, Manhattan y Canberra, siendo la primera de estas la que manifestó mayores niveles en el índice de Calinski y Harabasz.

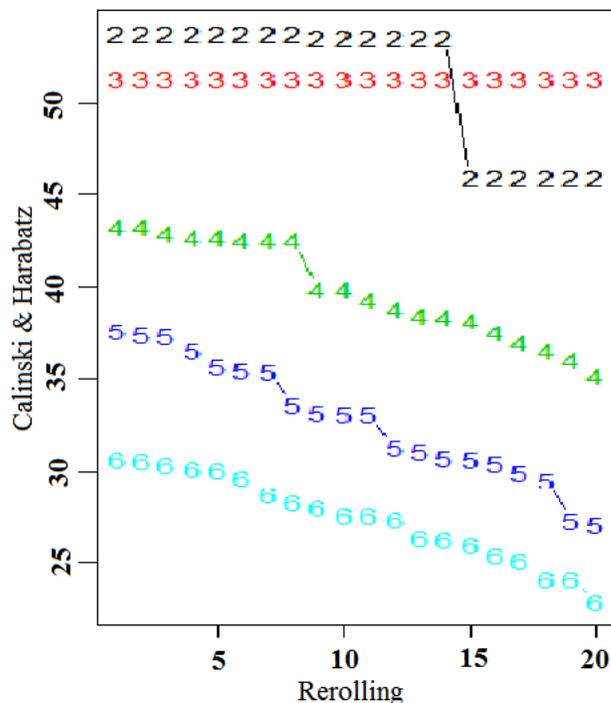
Tabla V.4

Promedio de índice de Calinski y Harabasz. (1974) según función distancia y número de grupos seleccionado.

Grupos	Euclidea	Maxima	Minkowski	Manhattan	Canberra
2	51	45	48	48	42
3	53	42	49	46	45
4	47	35	42	44	40
5	28	27	29	30	29
6	12	15	16	23	24

El número de grupos que sugiere este criterio considerando la distancia euclidiana es de 3. Este número de grupos maximiza también el criterio de optimización utilizando como disimilitud la distancia de Minkowski y Camberra.

Figura V.1
Índice de Calinski y Harabasz según re-lanzamiento del algoritmo. Distancia Euclídeana. Entre 2 y 6 agrupamientos.



Por su parte, la Figura V.I muestra los valores resultantes del índice de CH para evaluar la calidad del agrupamiento seleccionando entre 2 a 6 clústers para cada uno de los 20 sorteos del algoritmo. La clasificación obtenida a partir de “k-means longitudinal” para 2 grupos presenta inicialmente niveles promedio del índice mayores al resto de las asociaciones, sin embargo, la mayor estabilidad en los resultados se observa cuando definimos el clustering en 3 grupos.

Asimismo, se evaluaron otros cuatro criterios de clasificación provistos por el paquete utilizado: Una variante de Kryszczuk⁴¹ del índice Calinski y Harabasz, una variante de Christophe Genolini

⁴¹ Kryszczuk, K., & Hurley, P. (2010, April). Estimation of the number of clusters using multiple clustering validity indices. In International workshop on multiple classifier systems (pp. 114-123). Springer, Berlin, Heidelberg.

del índice Calinski y Harabasz, el criterio de Ray y Turi (1999)⁴² y el criterio de Davies y Bouldin (1979)⁴³. La figura D.6 del Anexo D muestra la estandarización en el rango 0 a 1 de los cinco criterios de selección de grupos en función del método de Calinski y Harabasz (1974). Se observa que, a excepción de la variante de Kryszczuk, los índices optimizan el agrupamiento longitudinal en 3 conglomerados.

V.5. ANÁLISIS DE LOS GRUPOS.

La información de los tres grupos conformados (Grupos: A, B y C) se resume en la Tabla V.5. En el grupo A fueron clasificadas el 63 por ciento de las cohortes de trabajadores (68), en el grupo B el 22.2% (24) y en el grupo C el 14,8% (16).

Las cohortes han sido conformadas por grupos de edad y sexo por lo que podemos caracterizar los grupos por las cohortes que quedaron clasificadas en cada uno. En grupo A incluye cohortes de varones que en el año base (1990) tenían edades comprendidas entre 20 y 59 años, y a mujeres que tenían edades entre 23 y 50 años. Este primer grupo se integra, en casi un 60 por ciento, por cohortes de asalariados varones. El grupo B se compone con cohortes de varones que tenían, en 1990, edades ubicadas entre 14 y 19 años; y con cohortes de mujeres con edades en los rangos de 14 a 22 años y de 51 a 59 años. Este segundo grupo es constituido en un 70 por ciento por asalariadas mujeres.

⁴² Ray, S., & Turi, R. H. (1999, December). Determination of number of clusters in k-means clustering and application in colour image segmentation. In Proceedings of the 4th international conference on advances in pattern recognition and digital techniques (pp. 137-143).

⁴³ Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. IEEE transactions on pattern analysis and machine intelligence, (2), 224-227.

Por su parte, en el grupo C, se incorporaron trabajadores varones y mujeres que tenían 6 años al año de inicio del estudio. Es decir, este grupo engloba las cohortes que se incorporan al mercado de trabajo⁴⁴ recién en el sexto año (1995) del periodo de análisis.

Tabla V.5

Cohortes contenidas en cada conglomerado. Cantidad total y según género.

GRUPO	COHORTES	Total		Varones		Mujeres	
		N	(%)	N	(%)	N	(%)
A	i_v20 - i_v59	68	63.0	40	58.8	28	41.2
	i_m23 - i_m50						
B	i_v14 - i_v19	24	22.2	6	25.0	18	75.0
	i_m14 - i_m22						
	i_m51 - i_m59						
C	i_v06 - i_v13	16	14.8	8	50.0	8	50.0
	i_m06 - i_m13						

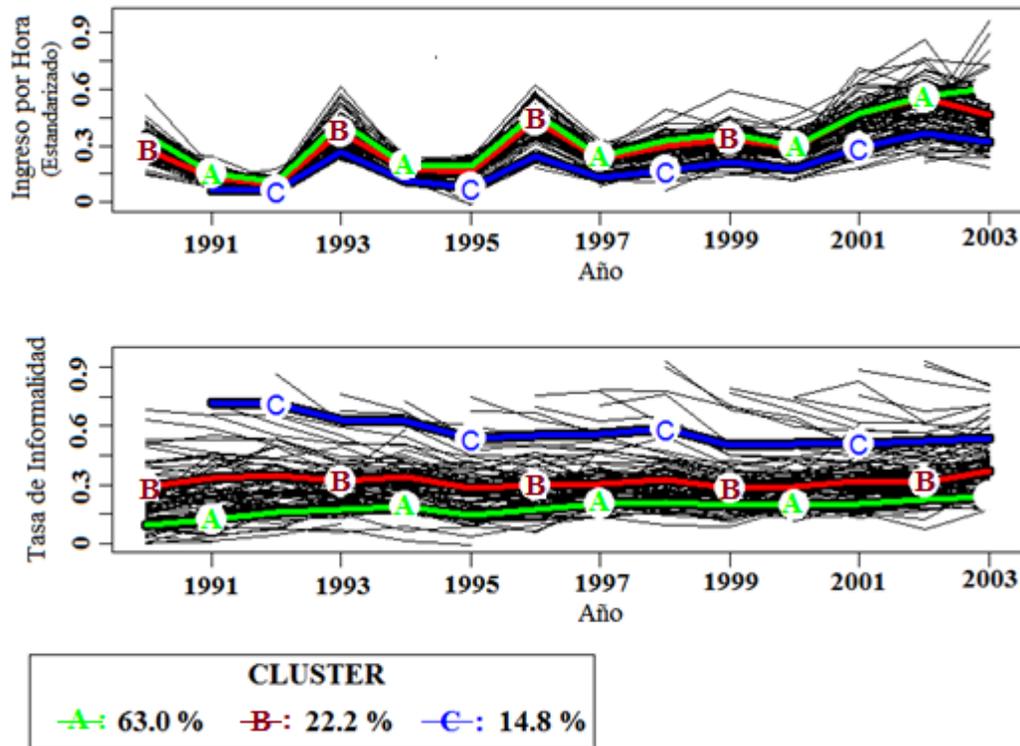
Un análisis de las trayectorias de cada una de las cohortes de trabajadores en los niveles de informalidad e ingreso laboral por hora se puede comparar con las trayectorias medias de cada uno de estos atributos en los tres clústers conformados⁴⁵ permite caracterizar la evolución de estas variables en cada grupo. (Figura V.2).

⁴⁴ Los trabajadores pueden haberse incorporado al mercado laboral en años previos a 1995. Sin embargo, dado los supuestos mencionados en la sección V.1, se consideró 18 años como el límite etario inferior para el análisis de las condiciones laborales.

⁴⁵ Tal como se mencionó previamente, el grupo C (azul), comienza su trayectoria media en 1995. En términos analíticos, se optó por medir la clasificación conjunta entre cohortes manteniendo los valores faltantes en los años previos dado que el paquete utilizado admite esta posibilidad conservando la calidad en sus resultados. Sin embargo, a fines ilustrativos, para obtener la figura V.2 se imputaron los datos de las cohortes en los años faltantes combinando estimaciones del modelo de pseudo panel y el método de imputación “*lineal interpolation*”. Los años estimados mediante modelación fueron el inicial (1990) y el intermedio (1993), mientras que, la técnica de interpolación completa la trayectoria trazando una línea entre los dos valores no perdidos que preceden y los siguen inmediatamente al valor perdido.

Figura V.2

Clasificación de KmL3D. Trayectorias medias de variables informalidad laboral e ingreso estandarizado para cada cluster. Distancia Euclidea. 3 conglomerados.

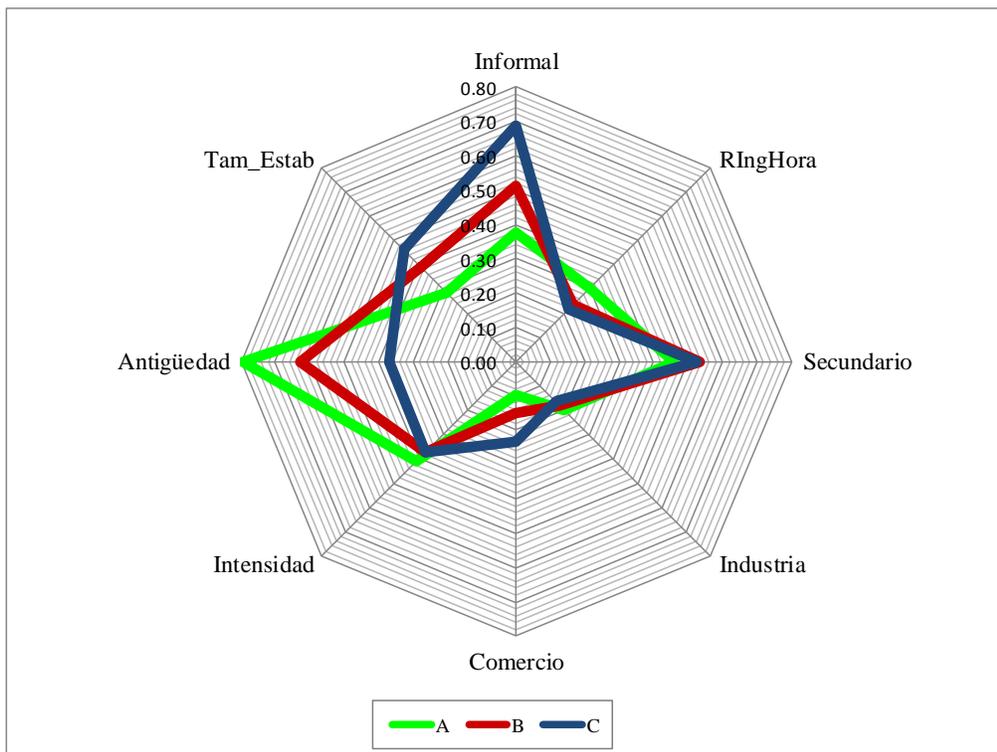


En la figura se destaca una importante correlación negativa entre el nivel medio de informalidad del conglomerado y el nivel de ingresos laborales medios que percibe cada grupo. El clúster A es el que admite simultáneamente menores niveles medios de informalidad laboral y mejores niveles de ingresos monetarios, mientras que, el clúster C presenta mayores niveles medios de empleo informal e ingresos medios menores. El clúster B presenta niveles intermedios en las trayectorias de ambas variables en relación a los demás grupos. Sin embargo, se advierte que la diferencia registrada entre los conglomerados B y C es mayor a la que se observa entre los conglomerados A y B, principalmente en la trayectoria de los ingresos laborales.

En sintonía con lo expresado, es admisible indicar la presencia de tres segmentos en el mercado de trabajo argentino (Figura V.3): El segmento de mejor condición ocupacional (A) es

conformado por cohortes de trabajadores, que en promedio, presentaron mayores niveles de antigüedad laboral, menor grado de subocupación horaria, proporción de empleo en microempresas inferior al 30 por ciento, menor participación relativa en actividades de comercio y mayores tasas de escolaridad completa.

Figura V.3
Dimensiones de los segmentos laborales.



Las cohortes incorporadas en el segmento de baja calidad laboral (C) se caracterizaron especialmente por poseer una menor antigüedad en los puestos de trabajo y presentar mayor participación relativa en microempresas y en tareas del sector comercial.

Por su parte, el segmento caracterizado por presentar condiciones intermedias de empleabilidad (B), es conformado principalmente por cohortes de mujeres y presenta en sus dimensiones

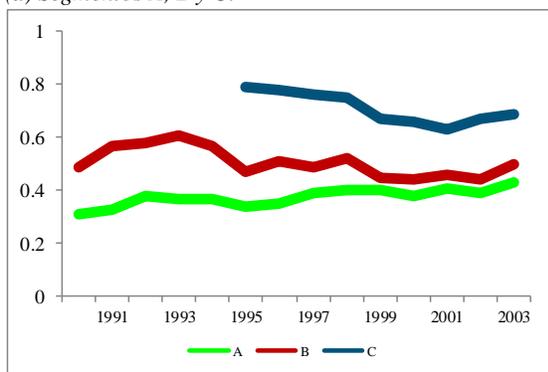
valores que se ubican en una posición media entre los otros dos grupos. Los atributos de las cohortes que lo componen se aproximan a las del segmento A en los niveles antigüedad y empleo en comercio, mientras que, se aproximan a las del segmento C en la tasa de escolaridad y el tamaño del establecimiento donde desarrollan sus tareas.

Las figuras V.4 y V.5 permiten visualizar la evolución temporal de los segmentos del mercado de trabajo, posibilitando inferir hipótesis respecto de su dinámica conjunta y del comportamiento interno en función del género de las cohortes.

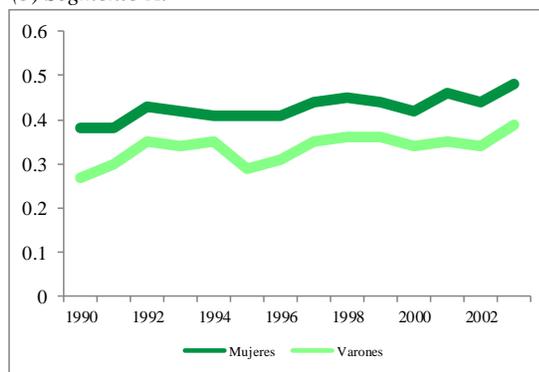
Figura V.4

Trayectorias de las tasas medias de informalidad laboral en los conglomerados desagregado según género.

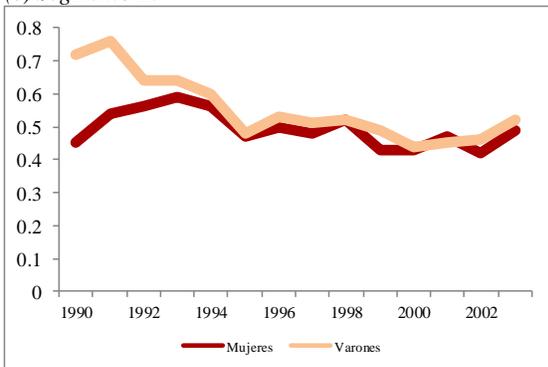
(a) Segmentos A, B y C.



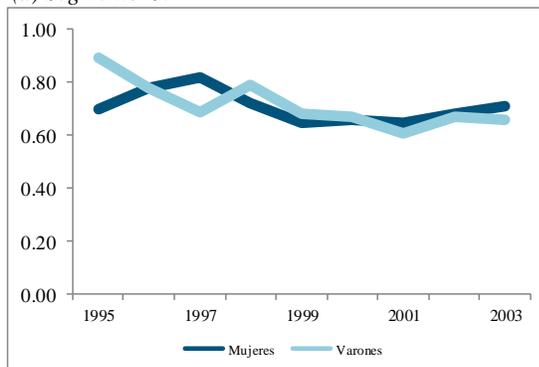
(b) Segmento A.



(c) Segmento B.



(d) Segmento C.



En relación a la dinámica de la informalidad, se observa que las trayectorias medias de los segmentos A y B tienden a converger para final de la década, debido a sucesivos incrementos en

la proporción de empleo no registrado en el primero de estos. Por su parte, el segmento C identificado por altos niveles de informalidad, exhibe una trayectoria media que no reduce temporalmente su brecha respecto del segmento intermedio B. Esto último, proporciona argumentos para suponer un posible deterioro en la calidad del empleo asalariado en Argentina durante este ciclo, principalmente a partir del año 1995.

Si examinamos la interseccionalidad de género, se observa que el comportamiento al interior de cada grupo es desigual:

En el Segmento A, caracterizado por mejores condiciones laborales, las cohortes de trabajadores mujeres presentan niveles medios de informalidad superiores al 10 por ciento en relación a sus pares varones. Esta brecha de género se mantiene constante durante todo el período.

En el segmento B, las tasas medias de informalidad entre ambos géneros presentan una diferencia inicial en favor de los varones, que logra converger a partir del año 1995. Las cohortes de varones incorporadas en este grupo presentan edades relativamente jóvenes. Por lo tanto, es probable que la diferencia inicial marcada por altos niveles de empleo no registrado en varones se explique por su proceso de ingreso al mercado de trabajo. Es decir, que esta disparidad tenga su origen en el “efecto edad” y no en el “efecto cohorte”. En el segmento C la penalización en las tasas medias de informalidad debido al género no se observa presente.

La evolución que presentan los segmentos según el ingreso percibido en la ocupación principal, plantea una dinámica temporal distinta a la que se observa en la informalidad laboral. La jerarquía entre los clústers mantiene el mismo ordenamiento: las cohortes del segmento A perciben en promedio mejores ingresos, las cohortes del segmento B se ubican en rangos de ingresos inferiores a este primer grupo y, a una distancia más que proporcional a la presente entre

los dos primeros, se ubica el segmento C con cohortes que perciben los menores ingresos laborales. Sin embargo, a diferencia de lo que sucede con la informalidad, los segmentos A y B no convergen temporalmente. Esto último, se exhibe en la Figura V.5 (a) en donde las series no reducen su brecha a final del periodo, sino que esta aumenta.

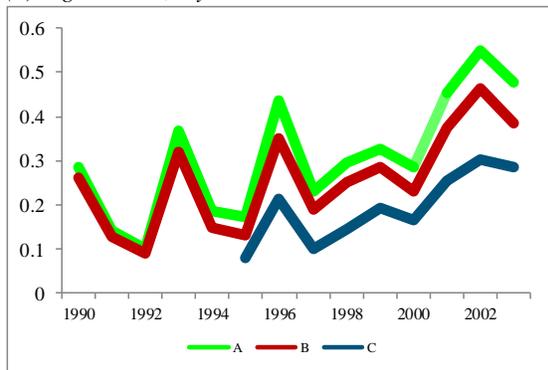
La interseccionalidad de género no parece estar presente en el comportamiento interior de cada grupo para el caso de los ingresos monetarios percibidos por hora de trabajo:

En los segmentos A y C, las series de los rangos medios de ingresos de las cohortes de trabajadores para cada género se superponen durante todo el periodo.

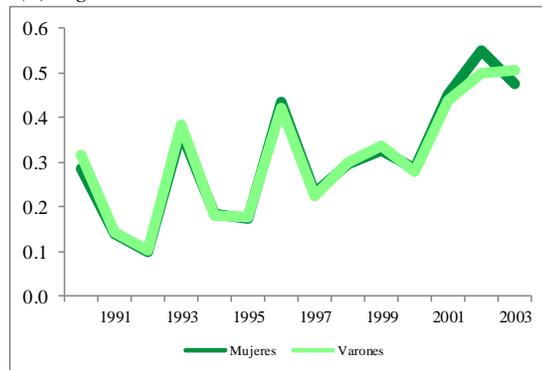
Figura V.5

Trayectorias de los ingresos laborales medios estandarizados en los conglomerados desagregado según género.

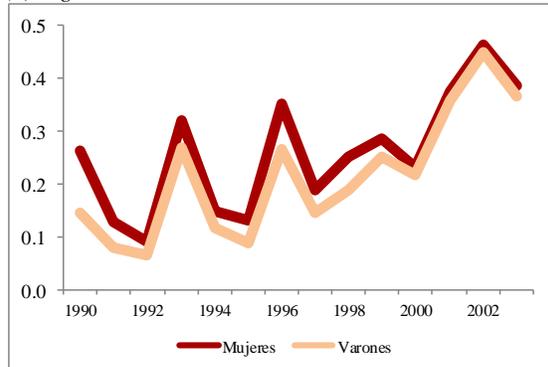
(a) Segmentos A, B y C.



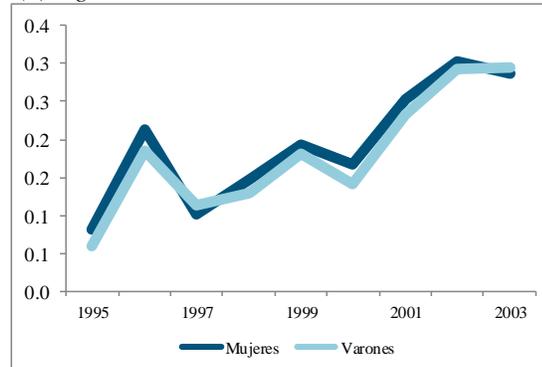
(b) Segmento A.



(c) Segmento B.



(d) Segmento C.



En el segmento B, se exhibe brecha de género por ingresos en favor de las cohortes de trabajadoras en los primeros años que se anula para el final de la década. No obstante, debido a la composición etaria de las cohortes que componen este grupo, es válido suponer que esta diferencia inicial, al igual que en el caso anterior, puede explicarse por motivos generacionales y no de género.

PARTE VI. CONCLUSIONES.

En el sexto capítulo se analizan los aportes realizados al campo del desarrollo económico y social, mediante la aplicación de las metodologías estadísticas implementadas. Posteriormente, se interpretan los principales resultados obtenidos y se ponen en discusión con dos de las principales teorías: la neoclásica y la teoría del mercado de trabajo segmentado. Finalmente, se presenta una síntesis de las principales conclusiones de la tesis y, en una segunda sección, se plantean potenciales líneas para continuar la investigación.

VI.1. CONCLUSIONES.

En Argentina, la economía informal, no representó niveles significativos hasta finalizar la primera mitad de la década del '70'. Las reformas económicas que se llevaron a cabo en los países latinoamericanos desde mediados de los 80 y su impacto sobre los mercados de trabajo provocaron un descenso en la calidad del empleo (Weller, 2003). Por lo que, durante la década del '90, el empleo informal en las áreas urbanas de América Latina creció aproximadamente de un 50% a un 58% (Gasparini y Tornarolli, 2007). En consecuencia, el estudio de la evolución de los perfiles del empleo informal a través de tiempo toma relevancia por las condiciones de “vulnerabilidad” a las que son expuestos estos trabajadores.

El enfoque de pseudo panel, permitió superar la falta de disponibilidad de datos temporales mediante la constitución de variables factores que definen cohortes de individuos que comparten características en común, el año de nacimiento y el género de los individuos, para este trabajo.

Existen algunos antecedentes de la aplicación de la metodología de pseudo panel en Argentina para el estudio de la pobreza (Casanova, 2008; Millón y García, 2018), la movilidad de ingresos (Navarro, 2006), y en el área específico del estudio de la informalidad, Arias & Escudero (2007), utilizan cohortes de paneles para comparar la evolución de las series de las tasas de informalidad y desempleo.

Sin embargo, en este trabajo el pseudo panel constituye una primera herramienta que proporciona la información necesaria para posteriormente definir perfiles de trabajadores, en función de similitudes presentadas en la evolución de la calidad de su empleo, utilizando un enfoque de k-medias longitudinal.

La combinación simultánea de estas dos metodologías resulta un aporte inédito en el área de desarrollo social y económico, que permite identificar relaciones entre las variables y las cohortes analizadas para extraer potenciales conclusiones en relación al fenómeno en estudio, cuando no se dispone de paneles completos.

En este sentido, los resultados obtenidos constituyen nuevos aportes en el estudio de la dinámica temporal del mercado de trabajo argentino desde un enfoque no ortodoxo, dado que refuerzan la evidencia de heterogeneidad en la estructura laboral, materializada en la existencia de segmentación socio-demográfica longitudinal.

La teoría neoclásica, que constituye una de las escuelas hegemónicas en ciencias económicas, interpreta al mercado de trabajo como un mercado parcial cualquiera dentro del sistema capitalista, que alcanza su equilibrio óptimo con un salario obtenido en la perfecta interacción entre la oferta y demanda de fuerza laboral. En consecuencia, lo entiendo como un mercado homogéneo en su estructura (Pries, 1997).

Sin embargo, las metodologías aplicadas en el análisis exploratorio resaltan la existencia de marcadas diferencias en las condiciones de empleo de los trabajadores asalariados argentinos, como así también, notorias variaciones temporales que se materializan en los ingresos percibidos y en el reconocimiento de sus derechos laborales.

La teoría del mercado de trabajo segmentado acepta la existencia de heterogeneidad de los puestos de trabajo entre mercados parciales diferentes y las restricciones de intercambiabilidad de la mano de obra para ocupar otros puestos. Este enfoque a las normas e instituciones sociales que estructuran las relaciones en el mercado laboral entre los trabajadores y los tipos de empleo (Sengenberger, 1988). La presente investigación, considera dicho marco teórico y extiende

conclusiones a que la heterogeneidad ocupacional admite también segmentación sociodemográfica en términos longitudinales.

La respuesta generalmente obtenida por parte de la Escuela Neoclásica, frente a los distanciamientos observados entre lo empírico y lo teórico, es atribuida a “imperfecciones” en el mercado causadas por “sobre regulaciones” inducidas por agentes externos, como el Estado o los sindicatos, que conducen a desequilibrios.

No obstante, a inicios de la década de los noventa, en el marco de las reformas estructurales, se introdujeron una serie de reformas en las regulaciones laborales que buscaron flexibilizar el mercado de trabajo con el objetivo de mejorar la competitividad internacional (Altimir y Beccaria, 1999). Este proceso de “desregulación” incluyó la disminución de los aportes patronales, la flexibilidad en la jornada máxima e instauró los contratos temporales y períodos de prueba.

La mayoría de estas medidas se promovieron en 1991 pero tuvieron vigencia efectiva a partir del año 1995, lo que coincide con el deterioro observado en las condiciones laborales de las cohortes de trabajadores asalariados que se incorporan al mercado de trabajo posterior a esta fecha y que son clasificados en un segmento de inferior calidad de empleo.

En síntesis:

- El análisis exploratorio aplicado en este trabajo, permitió identificar tres perfiles de cohortes de trabajadores informales bien definidos considerando las trayectorias temporales de sus niveles de informalidad e ingreso laboral.
- Los resultados admiten heterogeneidad en la estructura del mercado de trabajo argentino en el periodo analizado, observándose temporalmente segmentación sociodemográfica.

- Las interacciones de género y generación, exhibe la presencia de brecha en la calidad del empleo en favor de las cohortes masculinas y de las cohortes que se incorporaron al mercado de trabajo durante la primera etapa.
- La evidencia empírica indica, que la desregulación de las relaciones del mercado de trabajo argentino, no solo no tuvo efectos positivos en los niveles de contratación, sino que además, la dimensión del empleo no registrado aumento sustancialmente.
- El deterioró en la calidad del empleo en los grandes centros urbanos de Argentina debido al aumento medio en los niveles de informalidad se presentó tanto en periodos de crecimiento macroeconómico (1990, 1992-1994 y 1996-1998, 2000) como en períodos de contracción de la economía (1991, 1995, 1999 y 2001-2002).

VI.2. LINEAS DE TRABAJO.

La aplicación de estas metodologías para corroborar la presencia de segmentación longitudinal en las dos décadas siguientes, principalmente entre los años 2003 a 2008 en donde la región se distinguió por un crecimiento económico acelerado y sostenido con mejoras en las variables macroeconómicas, puede presentar un potencial estudio de interés en el área.

Las posibles líneas de investigación que pueden establecerse a partir de los resultados aportados en esta tesis, o de aquellos que solo fueron explorados solo parcialmente, son:

- Analizar la existencia de variaciones temporales en las ocupaciones típicas en cada género y generación.
- En caso de existir cambios temporales en la rama de ocupación empleada, examinar mediante las metodologías propuestas si la brecha según género y generación persisten.

- Evaluar el efecto que tuvieron las crisis macroeconómicas (1991, 1995, 1999 y 2001-2002) en los niveles generales de la calidad de empleo en cada segmento.
- Indagar mediante qué mecanismos las reformas aplicadas o cambio el esquema socio-productivo afectaron la calidad del empleo en cada segmento.
- Evaluar la evolución heterogeneidad en las condiciones de empleo debido a cambios en las características del establecimiento empleador.

PARTE VII. BIBLIOGRAFÍA.

- Altimir, O. & Beccaria, L. (1999). "El mercado de trabajo bajo el nuevo régimen económico en Argentina". Serie Reformas Económicas N° 28. Santiago de Chile: Comisión Económica para América Latina y el Caribe (CEPAL).
- Alzúa, M. L. (2008), Are informal workers secondary workers?: Evidence for Argentina, La Plata, Centro de Estudios Distributivos, Laborales y Sociales (CEDLAS), Universidad Nacional de La Plata, Documento de Trabajo N° 73.
- Antman, F. and McKenzie, D. (2007). "Earnings Mobility and Measurement Error: A Pseudo-Panel Approach". *Economic Development and Culture Change*. Volume 56, Number 1. October, 2007.
- Arellano, M. y Bover, O. (1990). "La econometría de datos de panel". *Investigaciones Económicas (Segunda época)*. Vol XIV, n.º 1 (1990), págs. 3-45.
- Arias, O., & Escudero, W. S. (2007). Assessing trends in informality in Argentina: A cohorts panel VAR approach. Mimeografía, Banco Mundial y CEDLAS.
- Barrett, J., Diggle, P.J., Henderson, R. and Taylor-Robinson, D. (2013). Joint modelling of repeated measurements and time-to-event outcomes: flexible model specification and exact likelihood inference. *Journal of the Royal Statistical Society, B* (submitted).
- Bay, G. (2001). Estimaciones indirectas de indicadores demográficos para áreas menores: situación en América Latina. *Notas de Población*.

- Beccaria, L. (2001), "Movilidad Laboral e Inestabilidad de Ingresos en Argentina", en Estudios del Trabajo N° 21, Buenos Aires, Asociación Argentina de Especialista en Estudios del Trabajo (ASET).
- Beccaria, L. y Groisman F. (2007). "Informalidad y pobreza en Argentina", en Anales de la Asociación Argentina de Economía Política, Buenos Aires, 2007.
- Beccaria, L. y Groisman F. (2008). "Informalidad y pobreza en Argentina: una relación compleja". Investigación Económica, vol. LXVII, 266, octubre-diciembre de 2008, pp. 135-169.
- Beccaria, L. y Groisman F. (2015). "Informalidad y segmentación del mercado laboral: el caso de la Argentina". Revista CEPAL, 117, 127-143.
- Berg, I. (1970). Education for Jobs; The Great Training Robbery.
- Bocchicchio, F., Lorenzetti, A., Pok, C., & Roggi, M. C. (2014). Las estrategias de manutención de los hogares de la Ciudad de Buenos Aires: una propuesta de medición y análisis. In III Jornadas Nacionales sobre estudios regionales y mercados de trabajo. Universidad Nacional de Jujuy (Facultad de Cs. Económicas y Unidad de Investigación en Comunicación, Cultura y Sociedad de la Facultad de Humanidades y Cs. Sociales) y Red SIMEL.
- Busso, M. (2005); " Le travail informel : entre théorie et expérience". Document LEST. Séminaire thèse; Juin 2005.
- Busso, M. (2007); "Travailleurs informels en Argentine: de la construction d'identités collectives à la constitution d'organisations ?". Sociology. Université de Provence - Aix-MarseilleI; Universidad de Buenos Aires, 2007.

- Busso, M. (2010) “Las crisis y el trabajo informal en la Argentina (O de cómo las crisis socio-económicas permean lugares de trabajo “atípicos”)”, Revista Atlántida N°2 (dic. 2010) págs. 125 a 138.
- Canavire-Bacarreza G., Urrego J. A., Saavedra F. (2017). "Informality and Mobility in the Labor Market: A pseudo-panel's approach". Revista Latinoamericana de Desarrollo Económico. N°.27. La Paz. Mayo, 2017.pp 57-75.
- Casanova, L. (2008). Trampas de pobreza en Argentina: evidencia empírica a partir de un pseudo panel (No. 64). Documento de Trabajo.
- Celeux, G., & Govaert, G. (1992). A classification EM algorithm for clustering and two stochastic versions. Computational statistics & Data analysis, 14(3), 315-332.
- Censo Nacional Argentina (1869). Primer censo de la República Argentina: Verificado en los días 15, 16 y 17 de setiembre de 1869.
- Censo Nacional Argentina (1895). Segundo Censo de la República Argentina, mayo 10 de 1895.
- Censo Nacional Argentina (1914). Tercer censo nacional levantado el 10 de junio de 1914.
- Censo Nacional Argentina (1947). IV Censo General de la Nación. 1947.
- Censo Nacional Argentina (1960). Censo nacional de población. Instituto Nacional de Estadística y Censos. 1960.
- Censo Nacional Argentina (1970). Censo Nacional de Población, Familias y Viviendas. 1970.
- Censo Nacional Argentina (1980). Censo Nacional de Población y Vivienda 1980. 1980.

- Censo Nacional Argentina (1991). Censo Nacional de Población y Vivienda 1991. Instituto Nacional de Estadística y Censos. 1991.
- Censo Nacional Argentina (2001). Censo Nacional de Población, Hogares y Viviendas 2001. . Instituto Nacional de Estadística y Censos. 2001.
- Censo Nacional Argentina (2010). Censo Nacional de Población, Hogares y Viviendas 2010. Instituto Nacional de Estadística y Censos & de empleados. 2010.
- Deaton, A. (1985) "Panel data from time series of cross-sections", *Journal of Econometrics*, vol. 30, no 1-2, p. 109–126.
- Diggle, P.J., Heagerty, P., Liang, K-Y and Zeger, S.L. (2002). *Analysis of Longitudinal Data* (second edition). Oxford: Oxford University Press.
- Everitt, B., Landau, S., & Leese, M. (2001). *Cluster Analysis*. A Hodder Arnold Publication. Willey, London.
- Feres, J. C. (1988). Las encuestas de hogares y la medición del ingreso en América Latina. *Estudios de economía*, 15(1), 325-340.
- Fernández, M. L. (2004). Cambios metodológicos en la encuesta permanente de hogares.
- Fernández-Huerga, E. (2010). La teoría de la segmentación del mercado de trabajo: enfoques, situación actual y perspectivas de futuro. *Investigación económica*, 69(273), 115-150.
- Ferré, M. B., Novell, M. I., & Moreso, J. P. (2009). Estudios de cohorte. *Fisioterapia*, 31(5), 218-223.

- Formichella, M. M., & Rojas, M. (2008). Un aporte a la evidencia empirica del efecto de la educacion sobre el empleo formal. *Estudios económicos*, 25(51), 49-72.
- Fuller, W. A. (2002). Regression estimation for survey samples. *Survey Methodology*, 28(1), 5-24.
- Fuller, W. A. (2009). *Measurement error models* (Vol. 305). John Wiley & Sons.
- Garre M., Cuadrado J.J., Sicilia M.A., Rodriguez D. & Rejas R. (2007). "Comparación de diferentes algoritmos de clustering en la estimación de coste en el desarrollo de software". *Revista Española de Innovación, Calidad e Ingeniería del Software*, Vol.3, No. 1, 2007.
- Gasparini, L. y Tornarolli, L. (2007). "Labor Informality in Latin American and the Caribbean: Patterns and Trends from Household Surveys Microdata". CEDLAS, Documento de Trabajo N° 46, La Plata, Febrero.
- Genolini C., Alacoque X., Sentenac M. and Arnaud C.(2015). "Kml and kml3d: R packages to Cluster Longitudinal Data". *Journal of Statistical Software*. Volume 65, Issue 4. May 2015.
- Gleser, L. J. (1991). *Measurement error models*. *Chemometrics and Intelligent Laboratory Systems*, 10(1-2), 45-57.
- Graña, J. M., & Lavopa, A. (2008). 15 años de EPH, una serie: Empalme entre sus versiones Puntual y Continua, 1992-2006 (No. 11). Documentos de Trabajo.
- Guillerm, M. (2015). "Les méthodes de pseudo-panel". Document de travail. Institut National de la Statistique et des Etudes Economiques. París. Fevrier, 2015.

- Guillerm, M. (2017). "Pseudo-panel methods and an example of application to Household Wealth data". *Economie et Statistique*, 2017, pp. 109-130.
- Hausman, J. A. (1978). "Specification Test in Econometrics", *Econometrica*, Vol. 46,Nº 6. Noviembre, 1978, pp-1251-1272.
- Hussmanns, R. (2004). "Measuring the informal economy: From employment in the informal sector to informal employment", documento de trabajo núm. 53 del Departamento de Integración de Políticas, OIT, 2004, Ginebra.
- Hussmans, R. (2004). "Defining and Measuring Informal Employment", OIT, Ginebra.
- Iglesias M. & Stimolo M.I. (2019). "Metodologías de análisis longitudinal para determinar los Perfiles de la Informalidad laboral en Gran Córdoba". XV Congreso Dr. Antonio Monteiro. Bahía Blanca. Junio, 2019.
- Iglesias M. & Stimolo M.I. (2019). "Perfiles de la informalidad laboral en Gran Córdoba". XLVII Coloquio Argentino de Estadística. Tucumán. Octubre, 2019.
- Instituto Nacional de Estadísticas y Censos (2018). "Informes Técnicos" vol. 2, nº119. Trabajo e Ingresos. Vol. 2, nº 4. Mercado de trabajo. Tasas e indicadores socioeconómicos (EPH). Primer trimestre de 2018.
- Isaki, C. T., & Fuller, W. A. (1982). Survey design under the regression superpopulation model. *Journal of the American Statistical Association*, 77(377), 89-96.
- Jensen, C.S & Snodgrass, R.T. "Semantics of Time-Varying Information, 2000, *Information Systems*, Vol. 21, Nº 4, 1996.

- Jensen, C.S & Soo, M.D. & Snodgrass, R.T. "Unifying Temporal Data Models, Information Systems, Vol. 19, N° 7,1994.
- Jiménez, M. (2010). "El empleo informal y la estructura del mercado laboral en las regiones argentinas". Revista de Estudios Regionales y Mercado de Trabajo N° 6. Año 2010. Págs. 139-179.
- Jiménez, M. (2011). "La Economía Informal y el Mercado Laboral en la Argentina: Un Análisis desde la Perspectiva del Trabajo Decente, Documento de Trabajo, No. 116, Universidad Nacional de La Plata, Centro de Estudios Distributivos, Laborales y Sociales (CEDLAS), La Plata.
- Kacef O. & Lopez-Monti, R. M. (2010). "América Latina del auge a la crisis: desafíos de política macroeconómica". Comissão Econômica para a América Latina e o Caribe – CEPAL, 2010.
- Kuznets, S. (2015). Inventive activity: Problems of definition and measurement. In The rate and direction of inventive activity (pp. 19-52). Princeton University Press.
- Lazcano-Ponce, E., Fernández, E., Salazar-Martínez, E., & Hernández-Avila, M. (2000). Estudios de cohorte. Metodología, sesgos y aplicación. Salud pública de México, 42, 230-241.
- Leys, C. (1973). Interpreting African Underdevelopment: Reflections on the ILO Report on Employment, Incomes and Equality in Kenya. The Royal African Society, Oxford University. African Affairs, Vol. 72, No. 289 (Oct., 1973), pp. 419-429.
- López, L. (2004). La Encuesta Permanente de Hogares. Población de Buenos Aires, 1, 38-44.
- Mario, A. & Garcia, A. (2013). Informalidad laboral, pobreza y regiones. Un análisis desde la coyuntura argentina. Revista Estudios Regionales y Mercado de Trabajo(9), pp. 107- 125.

- Marshall, A., & Perelman, L. (2006). Salario mínimo, mercado de trabajo y pobreza. Argentina (2003-2005). Documento presentado en la reunión técnica sobre informalidad, pobreza y salario mínimo, OIT-MTEySS, Buenos Aires, 26.
- Martine, G. (2005). O papel dos organismos internacionais na evolução dos estudos populacionais no Brasil: notas preliminares. *Revista Brasileira de Estudos de População*, 22, 257-275.
- Martine, G., Camarano, A. A., Fletcher, P. R., & Neupert, R. (2015). A PNAD: notas para uma avaliação. *Livros*, 281-310.
- Maurizio, R. (2012). "Labour informality in Latin America: the case of Argentina, Chile, Brazil and Peru", Manchester, University of Manchester, BWPI Working Paper 165, abril.
- Maurizio, R. (2014). "Labour formalization and declining inequality in Argentina and Brazil in the 2000s: A dynamic approach". International Labour Office, Geneva 22 (Switzerland).
- Medina, F., & Galván, M. (2007). Imputación de datos: teoría y práctica. Cepal.
- Meng Y., Brennan A., Purshouse R. & Otros (2014). "Estimation of own and cross price elasticities of alcohol demand in the UK. A pseudo-panel approach using the Living Costs and Food Survey 2001–2009". *Journal of Health Economics*. Volume 34, March 2014, Pages 96-103.
- Menni, A. M. (2004). Cómo rebuscársela: trabajo informal en tiempos de crisis. *Política y cultura*. nº 22. México., pp. 47-70.
- Messere, M. R., & Hozzowski, A. E. (2007). Encuesta permanente de hogares: actualización del diseño de sus muestras, 1974-2003.

- Millón, E., & García, B. R. (2018). Trampas de Pobreza: evidencia para las regiones de Argentina. *Atlantic Review of Economics (ARoEc)*, 1(2).

- Mispireta, M. L., Rosas, Á. M., Velásquez, J. E., Lescano, A. G., & Lanata, C. F. (2007). Transición nutricional en el Perú, 1991-2005. *Revista Peruana de Medicina Experimental y Salud Publica*, 24(2), 129-135.

- Monza, A. (1999). La evolución de la informalidad en el área metropolitana en los años noventa. Resultados e interrogantes. *Informalidad y exclusión social*, ed. J. Carpio, E. Klein, e I. Novacovsky, Fondo de Cultura Económica, Siempre, OIT, Buenos Aires.

- Navarro, A. I. (2006). Estimating Income Mobility in Argentina with pseudo-panel data. Preliminary Version Department of Economics, Universidad de San Andres and Universidad Austral.

- Neffa J. C., Perez P. & Otros. (2006). Macroeconomía, mercado de trabajo y grupo vulnerables: desafíos para el diseño de políticas públicas. Buenos Aires: ASOCIACIÓN TRABAJO Y SOCIEDAD. Centro de Estudios e Investigaciones Laborales. CIEL - PIETTE.

- Neffa, J. (2016). Informalidad, empleo no registrado y empleo precario. En M. P. S., *Derecho social y trabajo informal* (págs. pp 113-139). España: Comares.

- Neffa, J. C. (2009). Sector informal, precariedad, trabajo no registrado. 9º Congreso Nacional de Estudios del Trabajo. Buenos Aires: Asociación Argentina de Especialistas en Estudios del Trabajo.

- Novik, M. (2007). "Recuperando políticas públicas para enfrentar la informalidad laboral: el caso argentino 2003 – 2007". Material for the Tripartite Interregional Symposium on the Informal Economy: Enabling Transition to Formalization - Geneva. November 2007
- Organización Internacional del Trabajo (2002). Informe VI El trabajo decente y la economía informal 90ª reunión de la Conferencia Internacional del Trabajo, 2002.
- Organización Internacional del Trabajo (2002). Informe VI. El trabajo decente y la economía informal. Conferencia Internacional del Trabajo 90.a reunión 2002. Ginebra.
- Organización Internacional del Trabajo (2013). "La economía informal y el trabajo decente: una guía de recursos sobre políticas, apoyando la transición hacia la formalidad" / Oficina Internacional del Trabajo, Departamento de Política de Empleo. - Ginebra: OIT, 2013.
- Organización Internacional del Trabajo (2015). "Transitando de la informalidad a la formalidad". Serie Bibliografías Temáticas Digitales OIT, N°9. Lima: OIT.
- Organización Internacional del Trabajo (2018). "Mujeres y hombres en la economía informal: un panorama estadístico (tercera edición)". Organización Internacional del Trabajo – Ginebra: OIT, 2018.
- Paz, J. A. (2004). "Argentina: Dinámica de la Protección Social y el Mercado Laboral, 1995-2002". OIT, Santiago de Chile.
- Paz, J. A. (2015). Informalidad laboral y segmentación en la Argentina. Dimensión regional. En Caminos hacia la formalización laboral en Argentina. Colección de artículos sobre informalidad laboral. Buenos Aires (Argentina): Organización Internacional del Trabajo.

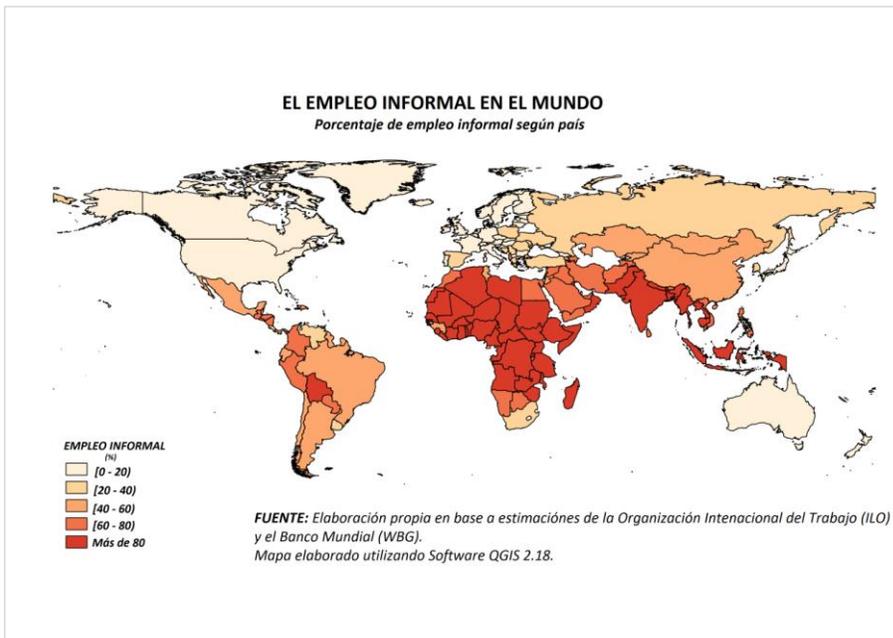
- Perera, J.M. (2006). "La movilidad de las rentas laborales en el mercado de trabajo uruguayo: Un enfoque de pseudo-panel". Documentos de Trabajo. CINVE., Montevideo. Junio, 2006.
- Perona, N., & Robin, S. (2000). Desocupación y vulnerabilidad: un estudio de caso en el barrio Parque Sur Rosario.
- Piselli, C. (2008). La Encuesta Permanente de Hogares. Instituto de Investigaciones Económicas. Universidad Nacional de Salta.
- Pries, L. (1997). Teoría sociológica del mercado de trabajo. Iztapalapa: Revista de Ciencias Sociales y Humanidades, (42), 71-98.
- Robin, S., & Crucella, C. (2002). Empleo y pobreza en el segundo aglomerado urbano del interior de la Argentina durante el segundo semestre de 1998.
- Robin, S., & Duran, P. (2005). Juventud, pobreza y exclusión en el Gran Rosario post devaluación. In Actas Séptimo Congreso Nacional de Estudios del Trabajo, Buenos Aires.
- Salvia, A. (2003). Mercados duales e informalidad en la Argentina: fragmentación y precarización de la estructura social del trabajo. In VI Congreso Nacional ASET. ASET.
- Salvia, A. (2003). Mercados segmentados en la Argentina: fragmentación y precarización de la estructura social del trabajo (1991-2002). Lavboratorio, 4(11-1), 5-11.
- Sengenberger, W. (1988). Introducción sobre la investigación del mercado de trabajo en la RFA. Instituciones y factores. Lecturas sobre el mercado de trabajo en la República Federal de Alemania (I), Mercado de trabajo, ocupación y desempleo.

- Singer & Jolly (1972). Employment, Incomes and Equality Lessons of the ILO Employment Strategy Mission to Kenya. Institute of Development Studies. University of Sussex.
- Weller, J. (2003). "Reformas económicas y situación del empleo en América Latina". Entre el trabajo y la política: las reformas de las políticas sociales argentinas en perspectiva comparada. Editorial Biblos, 2003.
- Wooldridge, J. M. (2006). Introductory econometrics: A modern approach, 3rd. New York: Thomson.
- Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT press.
- Zeger S, Liang K, Albert P (1988) Models for longitudinal data: a generalized estimating equation approach. Biometrics, 44:1049–1060.

PARTE VIII. ANEXOS.

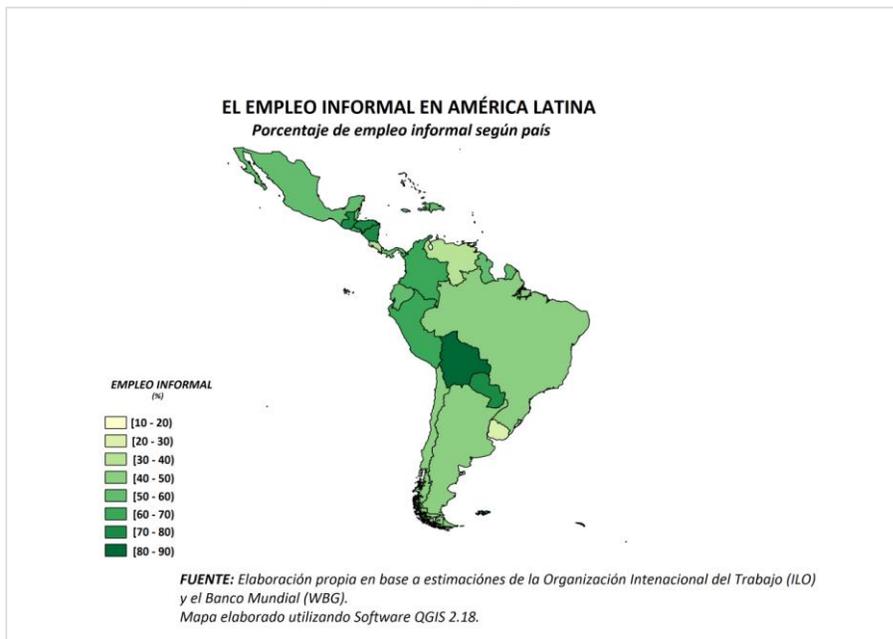
ANEXO I: INTRODUCCIÓN.

FIGURA I.1. Porcentaje de empleo informal según país.



Nota : El mapa indica según el gradiente de colores el porcentaje de empleo informal en cada país. Los datos fueron estimados por la Organización Internacional del Trabajo (OIT) para el año 2018.

FIGURA I.2. Porcentaje de empleo informal según país. América del Sur.



Nota : El mapa indica según el gradiente de colores el porcentaje de empleo informal en cada país de América del Sur. Los datos fueron estimados por la OIT para el año 2018.

ANEXO II: A. MATERIALES

Tabla A.1.

Principales programas oficiales de encuestas de hogares en países de América Latina.

País	Programa	Vigente
Argentina	Encuesta Permanente de Hogares Puntual (EPH)	No
	Encuesta Permanente de Hogares Continua (EPH)	Sí
Bahamas	Labour Force Survey (LFS)	Sí
Barbados	Continuous Labour Force Sample Survey (CLFSS)	Sí
Belice	Labour Force Survey (LFS)	Sí
	Encuesta Integrada de Hogares (EIH)	No
Bolivia	Encuesta Nacional de Empleo (ENE)	No
	Encuesta Continua de Hogares (ECH)	Sí
Brasil	Pesquisa Nacional por Amostra de Domicilio (PNAD)	Sí
Chile	Encuesta de Caracterización Socioeconómica Nacional (CASEN)	Sí
	Encuesta Nacional de Hogares - Fuerza de Trabajo (ENH-FT)	No
Colombia	Encuesta Continua de Hogares (ECH)	No
	Gran Encuesta Integrada de Hogares (GEIH)	Sí
	Encuesta de Hogares de Propósitos Múltiples (EHPM)	No
Costa Rica	Encuesta Nacional de Hogares (ENAHO)	Sí
	Encuesta Nacional de Fuerza de Trabajo (ENFT)	Sí
Ecuador	Encuesta Nacional de empleo, desempleo y subempleo (ENEMDU)	Sí
El Salvador	Encuesta de Hogares de Propósitos Múltiples (EHPM)	Sí
	Encuesta Nacional de Empleo e Ingresos (ENEI)	No
	Encuesta Nacional de Condiciones de Vida (ENCOVI)	No
Guatemala	Encuesta Nacional de Empleo e Ingresos (ENEI)	Sí
	Encuesta Permanente de Hogares de Propósitos Múltiples (EHPM)	Sí
Jamaica	Labour Force Survey (LFS)	Sí
México	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH)	Sí
	Encuesta de Hogares sobre medición de Niveles de Vida (EMNV)	Sí
Nicaragua	Encuesta Continua de Hogares (ECH)	Sí
	Encuesta de Hogares (EH)	No
	Encuesta de Hogares de Propósitos Múltiples (EHPM)	Sí
Panamá	Encuesta de Hogares - Mano de Obra (EHM)	No
	Encuesta Integrada de Hogares (EIH)	No
Paraguay	Encuesta Permanente de Hogares (EPH)	Sí
	Encuesta Nacional de Hogares (ENAHO)	Sí
Perú	Encuesta Nacional de Hogares (ENAHO)	Sí
Trinidad y Tobago	Continuous Sample Survey of Population (CSSP)	Sí
Uruguay	Encuesta Continua de Hogares (ECH)	Sí
Venezuela	Encuesta de Hogares por Muestreo (EHM)	Sí

FUENTE: Centro de Microdatos Armonizados de las Encuestas de Hogares de América Latina y el Caribe. Banco Interamericano de Desarrollo.

Tabla A.2.

Aglomerados urbanos relevados por Encuesta Permanente de Hogares en modalidad puntual según provincia y región.

AGLOMERADO	Provincia	Región	Población EPH (%)
Partidos de Buenos Aires	A.M.B.A.	Gran Buenos Aires	14.9
Ciudad Autónoma de Buenos Aires	A.M.B.A.	Gran Buenos Aires	39.9
San Salvador de Jujuy – Palpalá	Jujuy	Noroeste	1.1
Salta	Salta	Noroeste	1.9
Tucumán – Tafi Viejo	Tucumán	Noroeste	3.3
Gran Catamarca	Catamarca	Noroeste	0.6
La Rioja	La Rioja	Noroeste	0.5
Santiago del Estero – La Banda	Santiago del Estero	Noroeste	1.3
Formosa	Formosa	Nordeste	0.7
Resistencia	Chaco	Nordeste	1.5
Posadas	Misiones	Nordeste	1.1
Corrientes	Corrientes	Nordeste	1.3
Gran San Juan	San Juan	Cuyo	1.8
Gran Mendoza	Mendoza	Cuyo	3.9
San Luis - El Chorrillo	San Luis	Cuyo	0.6
Paraná	Entre Ríos	Pampeana	1.1
Concordia	Entre Ríos	Pampeana	0.6
Rosario	Santa Fé	Pampeana	5.6
Santa Fé y SantoTomé	Santa Fé	Pampeana	2
Córdoba	Córdoba	Pampeana	5.9
Río Cuarto - Las Higueras	Córdoba	Pampeana	0.7
Santa Rosa – Toay	La Pampa	Pampeana	0.4
Gran La Plata	Buenos Aires	Pampeana	3.2
Bahía Blanca	Buenos Aires	Pampeana	1.3
Mar del Plata-Batán	Buenos Aires	Pampeana	2.6
Neuquén – Plottier	Neuquén	Patagonia	0.9
Alto Valle (*)	Río Negro	Patagonia	1.0
Comodoro Rivadavia	Chubut	Patagonia	0.6
Río Gallegos	Santa Cruz	Patagonia	0.3
Tierra del Fuego	Tierra del Fuego	Patagonia	0.3
TOTAL			100.0

FUENTE: Elaboración propia en base a datos del "Manual de Consulta de Base Usuaría Ampliada (BUA)". Instituto Nacional de Estadísticas y Censos de Argentina (INDEC).

(*) Región urbana-rural.

Tabla A.3.*Principales características de los aglomerados urbanos en estudio.*

AGLOMERADO	Incorporación a EPH	Región	Población (2020)
Partidos de Buenos Aires	Octubre - 1972	Gran Buenos Aires	12397661
Ciudad Autónoma de Buenos Aires	Octubre - 1972	Gran Buenos Aires	2999749
Gran Córdoba	Mayo - 1978	Pampeana	1562487
Gran Rosario	Mayo - 1974	Pampeana	1320733
Gran Mendoza	Octubre - 1974	Cuyo	1024388
Gran Tucuman	Mayo - 1974	Noroeste	899931
Gran La Plata	Mayo - 1974	Pampeana	893844
TOTAL			21098793

FUENTE: Elaboración propia en base a datos del Instituto Nacional de Estadísticas y Censos de Argentina (INDEC).

Tabla A.4.*Cobertura poblacional de los aglomerados urbanos en estudio.*

GRUPO	Población	Cobertura (%)
Total aglomerados EPH	28537330	
Grandes Centros Urbanos	21098793	73.9
Total aglomerados de 500 mil y más	23446844	
Grandes Centros Urbanos	21098793	90.0

FUENTE: Elaboración propia en base a datos del Instituto Nacional de Estadísticas y Censos de Argentina (INDEC).

Tabla A.5 (a).*Composición administrativa y demográfica del Aglomerado urbano Gran Mendoza*

MUNICIPIO	Departamento	Población (2010)
Mendoza	Capital	114893
Guaymallén	Guaymallén	252618
Godoy Cruz	Godoy Cruz	191299
Las Heras	Las Heras	189067
Maipú	Maipú	106662
Luján de Cuyo	Luján de Cuyo	82615
Total		937154

FUENTE: Elaboración propia en base a datos de Censo Nacional de Población, Hogares y Viviendas 2010 y Departamento de Muestreo de la Encuesta Permanente de Hogares. Instituto Nacional de Estadísticas y Censos de Argentina (INDEC)

Tabla A.5 (b).

Composición administrativa y demográfica del Aglomerado urbano Partidos de Buenos Aires.

PARTIDO	Capital	Población (2010)
Almirante Brown	Adrogué	552902
San Miguel	San Miguel	276190
Avellaneda	Avellaneda	342677
Berazategui	Berazategui	324244
Cañuelas (b)	Cañuelas	51892
Escobar (b)	Belén de Escobar	213619
Esteban Echeverría	Monte Grande	300959
Ezeiza (a)	Ezeiza	163722
Florencio Varela	San Juan Bautista	426005
General Rodríguez (b)	General Rodríguez	87185
General San Martín	San Martín	414196
Hurlingham(a)	Hurlingham	181241
Ituzaingó (a)	Ituzaingó	167824
José C. Paz (a)	José C. Paz	265981
La Matanza	San Justo	1775816
Lanús	Lanús	459263
Lomas de Zamora	Lomas de Zamora	616279
Malvinas Argentinas (a)	Los Polvorines	322375
Marcos Paz (b)	Marcos Paz	54181
Merlo	Merlo	528494
Moreno	Moreno	452505
Morón	Morón	321109
Pilar (b)	Pilar	299077
Presidente Perón (b)	Guernica	81141
Quilmes	Quilmes	582943
San Fernando	San Fernando	163240
San Isidro	San Isidro	292878
San Vicente (b)	San Vicente	59478
Tigre	Tigre	376381
Tres de Febrero	Caseros	340071
Vicente López	Olivos	269420
Total		10763288

FUENTE: Elaboración propia en base a datos de Censo Nacional de Población, Hogares y Viviendas 2010 y Departamento de Muestreo de la Encuesta Permanente de Hogares. Instituto Nacional de Estadísticas y Censos de Argentina (INDEC)

NOTAS: (a) El partido fue creado en el año 1994.

(b) Incorporadas a partir de la onda de Mayo - 1998.

Tabla A.5 (c).

Composición administrativa y demográfica del Aglomerado urbano Gran Tucumán.

MUNICIPIO	Departamento	Población (2010)
San Miguel	Capital	548866
Banda del Río Salí (a)	Cruz Alta	63226
Yerba Buena (b)	Yerba Buena	59095
Las Talitas (c)	Tafi Viejo	52960
Alderetes (d)	Cruz Alta	41947
Diagonal Norte (f)	Tafi Viejo	11494
El Manantial (e)	Lules	16739
Total		794327

FUENTE: Elaboración propia en base a datos de Censo Nacional de Población, Hogares y Viviendas 2010 y Departamento de Muestreo de la Encuesta Permanente de Hogares. Instituto Nacional de Estadísticas y Censos de Argentina (INDEC)

NOTAS: (a) Incluye Lastenia

(b) Incluye Marcos Paz, Ex Ingenio San José y Barrio San José III.

(c) Incluye Villa M. Moreno y El Colmenar

(d) Incluye El Corte y Los Gutiérrez

(e) Incluye Barrio Araujo y Barrio San Felipe

(f) Incluye Luz y Fuerza, Los Pocitos y Villa Nueva Italia

Tabla A.5 (d).

Composición administrativa y demográfica del Aglomerado urbano Gran La Plata.

MUNICIPIO	Partido	Población (2010)
La Plata	La Plata	654324
Berisso	Berisso	88470
Ensenada	Ensenada	56729
Total		799523

FUENTE: Elaboración propia en base a datos de Censo Nacional de Población, Hogares y Viviendas 2010 y Departamento de Muestreo de la Encuesta Permanente de Hogares. Instituto Nacional de Estadísticas y Censos de Argentina (INDEC)

Tabla A.5 (e).

Composición administrativa y demográfica del Aglomerado urbano Gran Córdoba.

MUNICIPIO	Departamento	Población (2010)
Córdoba	Capital	1329604
La Calera	Colón	31972
Dumesnil	Colón	27738
Villa Allende	Colón	19688
Río Ceballos	Colón	18086
Unquillo	Colón	9612
Salsipuedes	Colón	
El Pueblito	Colón	
Villa El Fachinal	Colón	
Parque Norte	Colón	5710
Guiñazú Norte	Colón	
Mendiolaza	Colón	10271
Saldán	Colón	10605
La Granja	Colón	2456
Agua de Oro	Colón	1918
El Manzano	Colón	892
Canteras El Sauce	Colón	290
Total		1468842

FUENTE: Elaboración propia en base a datos de Censo Nacional de Población, Hogares y Viviendas 2010 y Departamento de Muestreo de la Encuesta Permanente de Hogares. Instituto Nacional de Estadísticas y Censos de Argentina (INDEC)

NOTAS: En la modalidad Puntual del programa EPH solo se incluyeron las localidades de Córdoba, Villa Allende, Saldán, Guiñazú Norte. A partir del Censo 2001, INDEC, incorpora el resto de las localidades a la constitución del aglomerado que pasaría a formar parte del programa EPH a partir de su modalidad continua (2003).

Tabla A.5 (f).

Composición administrativa y demográfica del Aglomerado urbano Ciudad de Buenos Aires

	Población (2010)
Ciudad Autónoma de Buenos Aires	2890151

FUENTE: Elaboración propia en base a datos de Censo Nacional de Población, Hogares y Viviendas 2010 y Departamento de Muestreo de la Encuesta Permanente de Hogares. Instituto Nacional de Estadísticas y Censos de Argentina (INDEC)

Tabla A.5 (g).

Composición administrativa y demográfica del Aglomerado urbano Gran Rosario

MUNICIPIO	Departamento	Población (2010)
Rosario	Rosario	948312
Va. Gdor. Galvez	Rosario	80769
G. Baigorria	Rosario	36994
Perez	Rosario	26448
Funes	Rosario	23281
San Lorenzo	San Lorenzo	45958
Cap. Bermudez	San Lorenzo	29218
F.L.Beltran	San Lorenzo	15176
Pto. San Martin	San Lorenzo	13243
Roldan	San Lorenzo	14113
Soldini (a)	Rosario	2577
Total		1236089

FUENTE: Elaboración propia en base a datos de Censo Nacional de Población, Hogares y Viviendas 2010 y Departamento de Muestreo de la Encuesta Permanente de Hogares. Instituto Nacional de Estadísticas y Censos de Argentina (INDEC)

NOTAS: (a) Incorporadas a partir de la onda de Mayo - 1998.

Tabla A.6.

Evolución demográfica de los aglomerados urbanos en estudio según registros censales y proyecciones oficiales. 1869 - 2020.

Aglomerado	AÑO DEL REGISTRO										
	1869 (a)	1895 (b)	1914 (c)	1947 (d)	1960 (e)	1970 (f)	1980 (g)	1991 (h)	2001 (i)	2010 (j)	2020 (k)
Gran La Plata (1)		45,410	137,413	266,000	404,129	485,939	564,750	642,979	694,253	799,523	893,844
Gran Mendoza	8,124	28,602	58,790	212,000	330,727	477,810	612,777	773,113	848,660	937,154	1,024,388
Gran Tucumán	17,438	34,305	92,284	203,555	297,305	366,392	498,579	622,324	738,479	794,327	899,931
Gran Rosario	23,169	91,669	245,199	485,000	669,173	806,942	956,761	1,118,905	1,161,188	1,236,089	1,320,733
Gran Córdoba	34,458	47,609	121,982	380,000	591,563	792,925	1,004,929	1,208,554	1,368,301	1,468,842	1,562,487
C.A.B.A.	187,787	663,854	1,575,814	2,981,043	2,966,634	2,972,453	2,922,829	2,965,403	2,776,138	2,890,151	2,999,749
Partidos B.A.		117,763	458,217	1,741,338	3,772,411	5,380,158	6,843,201	7,969,324	8,684,437	9,916,715	12,397,661
Gran B.A.	187,347	781,617	2,034,031	4,722,381	6,739,045	8,352,611	9,766,030	10,934,727	11,460,575	12,806,866	15,397,410

FUENTE: Elaboración propia en base a datos de Censos Nacionales de Población, Hogares y Viviendas de Argentina.

NOTA S: (1) La ciudad de La Plata fue planificada y construida en el año 1880 y fundada oficialmente el día 19 de noviembre de 1882.

(2) Las localidades satélites incorporan información poblacional a la ciudad central mediante a partir de los años en que conforman el aglomerado urbano según el criterio definido previamente.

(a) Primer censo de la República Argentina: Verificado en los días 15, 16 y 17 de setiembre de 1869.

(b) Segundo Censo de la República Argentina, mayo 10 de 1895.

(c) Tercer censo nacional levantado el 10 de junio de 1914.

(d) IV Censo General de la Nación. 1947.

(e) Censo nacional de población. Instituto Nacional de Estadística y Censos. 1960.

(f) Censo Nacional de Población, Familias y Viviendas. 1970.

(g) Censo Nacional de Población y Vivienda 1980. 1980.

(h) Censo Nacional de Población y Vivienda 1991. Instituto Nacional de Estadística y Censos. 1991.

(i) Censo Nacional de Población, Hogares y Viviendas 2001. Instituto Nacional de Estadística y Censos. 2001.

(j) Censo Nacional de Población, Hogares y Viviendas 2010. Instituto Nacional de Estadística y Censos & de empleados. 2010.

(k) Proyecciones de Población utilizadas en Encuesta Permanente de Hogares en el tercer trimestre de 2020. Instituto Nacional de Estadística y Censos.

*La ciudad de La Plata fue planificada y construida en el año 1880 y fundada oficialmente el día 19 de noviembre de 1882.

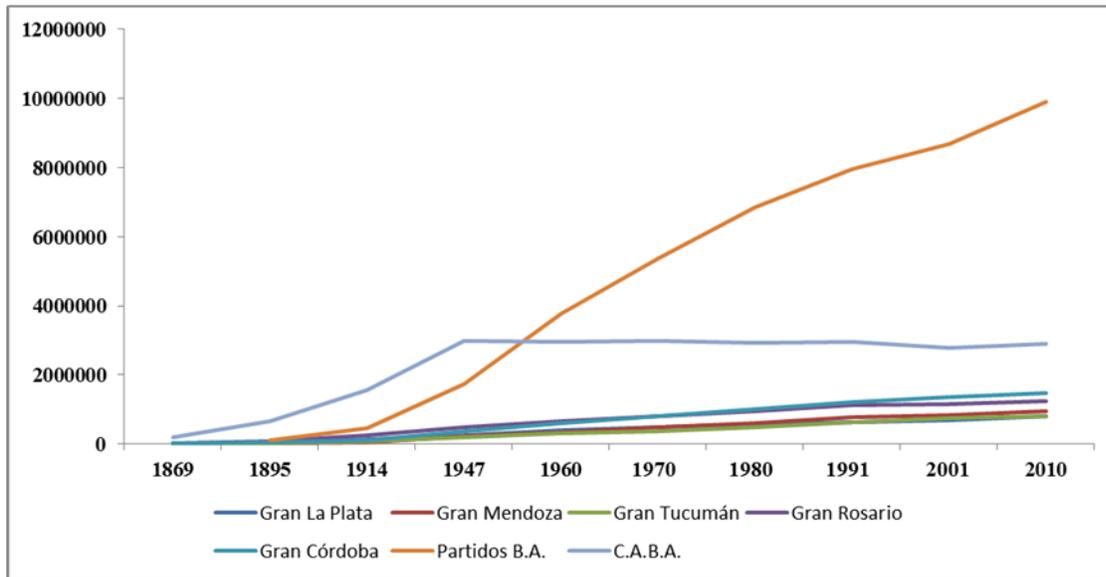
específicamente para que sirviera como capital de la provincia después de que la ciudad de Buenos Aires fuera declarada como distrito federal

Tabla A.7.

Diccionario de variables de cohortes de trabajadores asalariados en base de datos de trabajo.

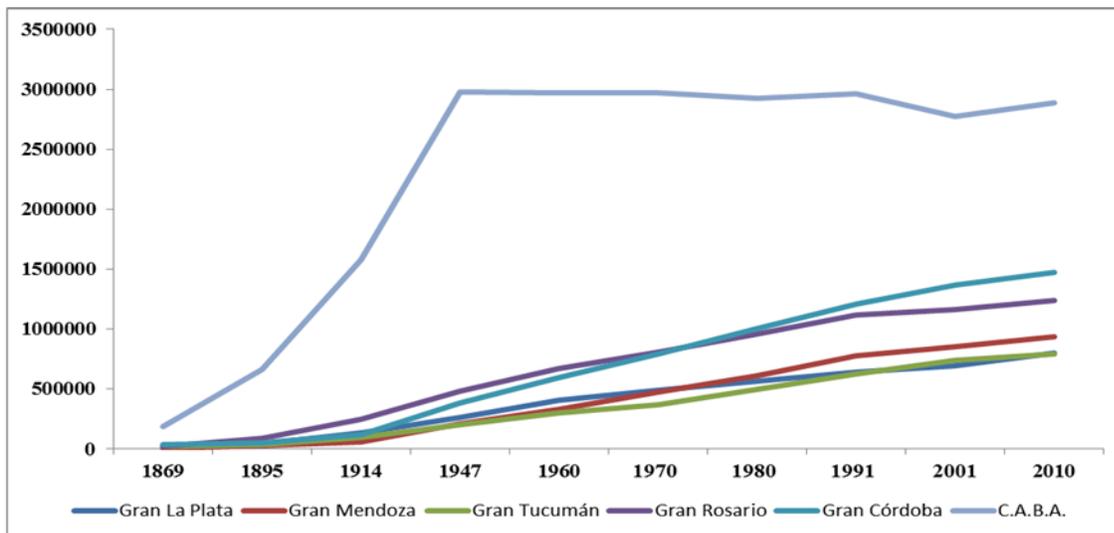
VARIABLE	DESCRIPCIÓN
Informal	Tasa informalidad en la cohorte.
Primario	Proporción de trabajadores con nivel primario completo en la cohorte.
Secundario	Proporción de trabajadores con nivel de secundario completo en la cohorte.
Industria	Proporción de trabajadores de la cohorte que desarrollan su ocupación principal en el sector industria.
Comercio	Proporción de trabajadores de la cohorte que desarrollan su ocupación principal en el sector comercio.
Construcción	Proporción de trabajadores de la cohorte que desarrollan su ocupación principal en el sector construcción.
Intensidad	Proporción de trabajadores de la cohorte con subocupación horaria en su ocupación principal.
Antigüedad	Proporción de trabajadores de la cohorte con antigüedad mayor a un año en su ocupación principal.
Tam_Estab	Proporción de trabajadores de la cohorte que desarrollan su ocupación principal en establecimientos de menos de 5 empleados.

FIGURA A.I. Evolución poblacional de los principales centros urbanos de Argentina de 1869 a 2020.



Nota: El gráfico representa la evolución demográfica de los aglomerados urbanos en estudio según registros censales y proyecciones oficiales entre los 1869 y 2020.

FIGURA A.II. Evolución poblacional de los principales centros urbanos de Argentina de 1869 a 2020. (excluyendo el aglomerado Partidos de Buenos Aires)



Nota: El gráfico representa la evolución demográfica de los aglomerados urbanos en estudio según registros censales y proyecciones oficiales entre los 1869 y 2020.

ANEXO III: B. ANALISIS EXPLORATORIO.

Tabla B.1.

Medidas de forma en características de los asalariados en el período 1990-2003.

Variable	Asimetría (1)	Curtosis (2)
Informal	0.91	1.35
Primario	-1.90	6.21
Secundario	0.00	-0.37
Industria	0.22	-0.53
Comercio	0.79	0.91
Construccion	1.11	0.59
Intensidad	0.85	1.09
Antigüedad	-1.42	2.19
Tam_Estab	1.03	1.67
IngHora	3.40	10.71
RIngHora	0.74	0.61
Edad	0.00	-1.20
Edad2	0.39	-1.08
EdadR	-0.22	-1.12

FUENTE: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual.

NOTAS: (1) Método "excess". Bilateral.

(2) Método "moment". Bilateral.

Tabla B.2.

Edades de cohortes de trabajadores asalariados con empleo informal extremadamente alto según año. Grandes centros urbanos de Argentina. Período 1990-2003.

Edad	Año														Edad	Total	%
	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003			
18	2	2	1	2	2	1	2	2	2	2	2	2	2	2	18	26	51.0
19	-	1	-	1	-	-	-	1	-	-	1	1	1	2	19	8	15.7
20	-	-	-	-	-	-	-	-	-	-	-	-	-	2	20	2	3.9
21	-	-	-	-	-	-	-	-	-	-	-	-	-	1	21	1	2.0
46	-	-	-	-	-	-	-	-	-	-	-	-	-	1	46	1	2.0
61	-	-	-	1	-	1	-	-	-	-	-	-	-	-	61	2	3.9
62	-	1	1	-	-	-	-	-	1	-	-	-	-	-	62	3	5.9
63	1	-	-	1	-	1	-	-	-	-	-	-	-	1	63	4	7.8
65	1	-	1	-	-	-	-	-	-	-	1	1	-	-	65	4	7.8
Total	4	4	3	5	2	3	2	3	3	2	4	4	3	9	Total	51	100.0

Fuente: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual.

Tabla B.3.

Género de cohortes de trabajadores asalariados con empleo informal extremadamente alto según año. Grandes centros urbanos de Argentina. Período 1990-2003.

Género	Año														Género Total	%	
	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003			
Mujer	3	2	3	4	1	2	1	2	2	1	3	3	2	6	Mujer	35	68.6
Varón	1	2	0	1	1	1	1	1	1	1	1	1	1	3	Varón	16	31.4
Total	4	4	3	5	2	3	2	3	3	2	4	4	3	9	Total	51	100.0

Fuente: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual.

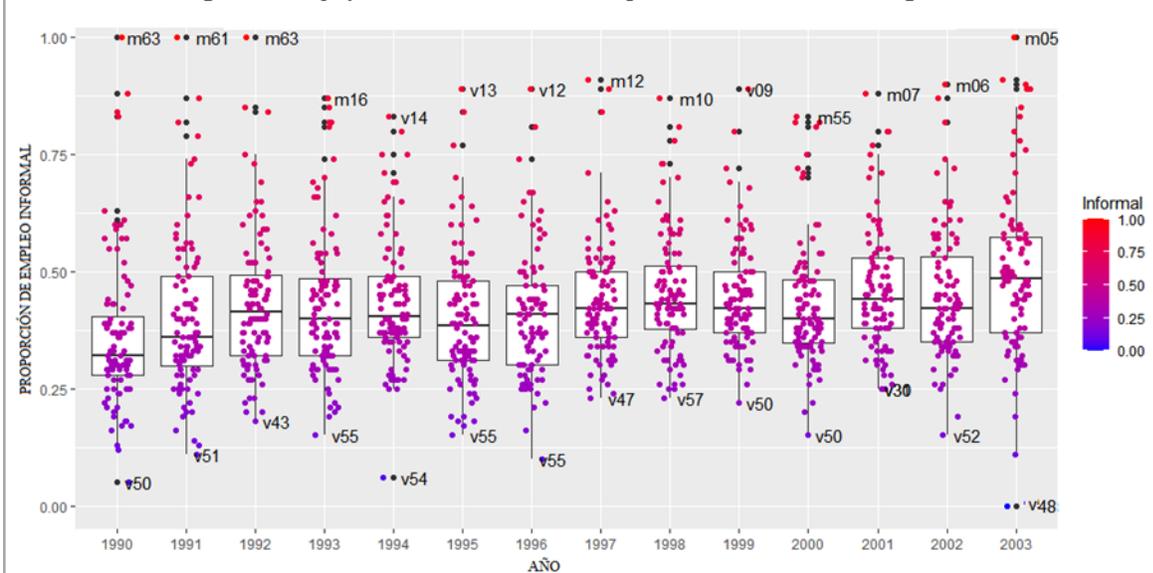
Tabla B.4.

Código de cohortes de trabajadores asalariados con empleo informal extremadamente alto según año. Grandes centros urbanos de Argentina. Período 1990-2003.

1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
m18	m17	m16	m15	m14	m56	m12	m11	m10	m09	m08	m07	m06	m05
m63	m61	m60	m16	v14	m58	v12	m12	m54	v09	m09	m08	m07	m06
m65	v17	m63	m58	-	v13	-	v11	v11	-	m55	m54	v06	m07
v18	v18	-	m60	-	-	-	-	-	-	v08	v07	-	m08
-	-	-	v15	-	-	-	-	-	-	-	-	-	m33
-	-	-	-	-	-	-	-	-	-	-	-	-	m50
-	-	-	-	-	-	-	-	-	-	-	-	-	v05
-	-	-	-	-	-	-	-	-	-	-	-	-	v06
-	-	-	-	-	-	-	-	-	-	-	-	-	v07

Fuente: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual.

FIGURA B.1. Diagrama de caja y brazos: Distribución de empleo informal en cohortes según año.



NOTA: La figura muestra la distribución de empleo informal en las cohortes de trabajadores asalariados en cada uno de los años en estudio. Además se indican la intensidad del empleo informal en paleta azulgrana y el código de las cohortes con niveles extremadamente atípicos.

Tabla B.5.*Medidas descriptivas. Informalidad laboral según año.*

	Mín.	1Q	Mediana	Media	3Q	Máx.
1990	0.0500	0.2800	0.3200	0.3619	0.4050	1.0000
1991	0.1100	0.2975	0.3600	0.4004	0.4900	1.0000
1992	0.1800	0.3200	0.4150	0.4317	0.4925	1.0000
1993	0.1500	0.3200	0.4000	0.4315	0.4850	0.8700
1994	0.060	0.360	0.405	0.438	0.490	0.830
1995	0.1500	0.3100	0.3850	0.3993	0.4800	0.8900
1996	0.1000	0.3000	0.4100	0.4089	0.4700	0.8900
1997	0.2300	0.3600	0.4200	0.4383	0.5000	0.9100
1998	0.2300	0.3775	0.4300	0.4534	0.5125	0.8700
1999	0.2200	0.3700	0.4200	0.4451	0.5000	0.8900
2000	0.1500	0.3475	0.4000	0.4319	0.4825	0.8300
2001	0.250	0.380	0.440	0.461	0.530	0.880
2002	0.1500	0.3500	0.4200	0.4470	0.5325	0.9000
2003	0.0000	0.3700	0.4850	0.4980	0.5725	1.0000

FUENTE: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual.

Tabla B.6.

Pruebas de normalidad en características de los asalariados en el período 1990-2003.
Bilateral.

Variable	SW (1)		KS (2)		LF (3)		AD (4)	
	Estadístico	p-valor	Estadístico	p-valor	Estadístico	p-valor	Estadístico	p-valor
Informal	0.95	0.0098	0.09	0.0097	0.09	0.0047	17.00	0.0037
Primario	0.84	0.0025	0.14	0.0000	0.14	0.0004	46.28	0.0004
Secundario	0.99	0.0526	0.04	0.0491	0.04	0.0173	1.65	0.0309
Industria	0.99	0.0151	0.06	0.0036	0.06	0.0073	4.40	0.0065
Comercio	0.96	0.0077	0.08	0.0011	0.08	0.0084	9.91	0.0059
Construccion	0.81	0.0087	0.65	0.0023	0.23	0.0095	92.69	0.0037
Intensidad	0.93	0.0034	0.65	0.0045	0.12	0.0011	23.66	0.0008
Antigüedad	0.88	0.0015	0.58	0.0024	0.15	0.0012	46.06	0.0089
Tam_Estab	0.95	0.0096	0.09	0.0022	0.09	0.0018	16.57	0.0091
IngHora	0.37	0.0001	0.48	0.0025	0.48	0.0088	381.66	0.0004
RIngHora	0.96	0.0086	0.06	0.0063	0.06	0.0034	10.73	0.0073
Edad	0.95	0.0085	0.07	0.0096	0.07	0.0094	15.27	0.0009
Edad2	0.93	0.0093	0.10	0.0010	0.10	0.0011	25.60	0.0090
EdadR	0.95	0.0095	0.08	0.0028	0.08	0.0014	16.25	0.0097

FUENTE: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual.

NOTAS: (1) Prueba "Shapiro-Wilk normality"

(2) Prueba "Kolmogorov-Smirnov normality"

(3) Prueba "Lilliefors normality"

(4) Prueba "Anderson-Darling normality"

Tabla B.7.

Distribución porcentual de personas ocupados según categoría ocupacional en los grandes centros urbanos de Argentina en el período 1990-2003.

Año	Cuenta Propia (%)	Asalariados (%)	Patrón (%)	Trab. Fliar Sin Remun. (%)
1990	25.6	68.7	4.4	1.3
1991	25.2	69.0	4.9	1.0
1992	24.2	69.8	4.7	1.3
1993	25.4	67.6	5.2	1.8
1994	25.1	68.4	5.2	1.4
1995	22.3	71.2	5.1	1.4
1996	22.7	70.4	5.1	1.7
1997	22.1	71.6	4.7	1.7
1998	22.2	71.8	4.7	1.3
1999	21.7	72.5	4.7	1.2
2000	20.8	73.0	5.2	1.0
2001	22.1	72.3	4.4	1.2
2002	24.2	71.1	3.5	1.1
2003	23.7	71.4	3.8	1.1

Fuente: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual.

Tabla B.8.

Porcentaje de trabajadores asalariados ocupados en empleos informales y porcentaje de trabajadores desocupados en los grandes centros urbanos de Argentina en el período 1990-2003.

Año	Tasa de Informalidad	Tasa de Desocupación
1990	35.6	7.5
1991	40.7	4.7
1992	44.5	5.7
1993	44.6	8.9
1994	44.3	9.2
1995	40.4	17.0
1996	42.0	16.3
1997	45.0	14.9
1998	46.5	12.2
1999	45.4	14.0
2000	44.3	14.7
2001	45.6	15.4
2002	45.3	20.3
2003	50.6	13.2

Fuente: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual.

FIGURA B.2. Histograma: Antigüedad laboral, Empleo en Comercio, Empleo en Construcción e ingresos laborales según género.

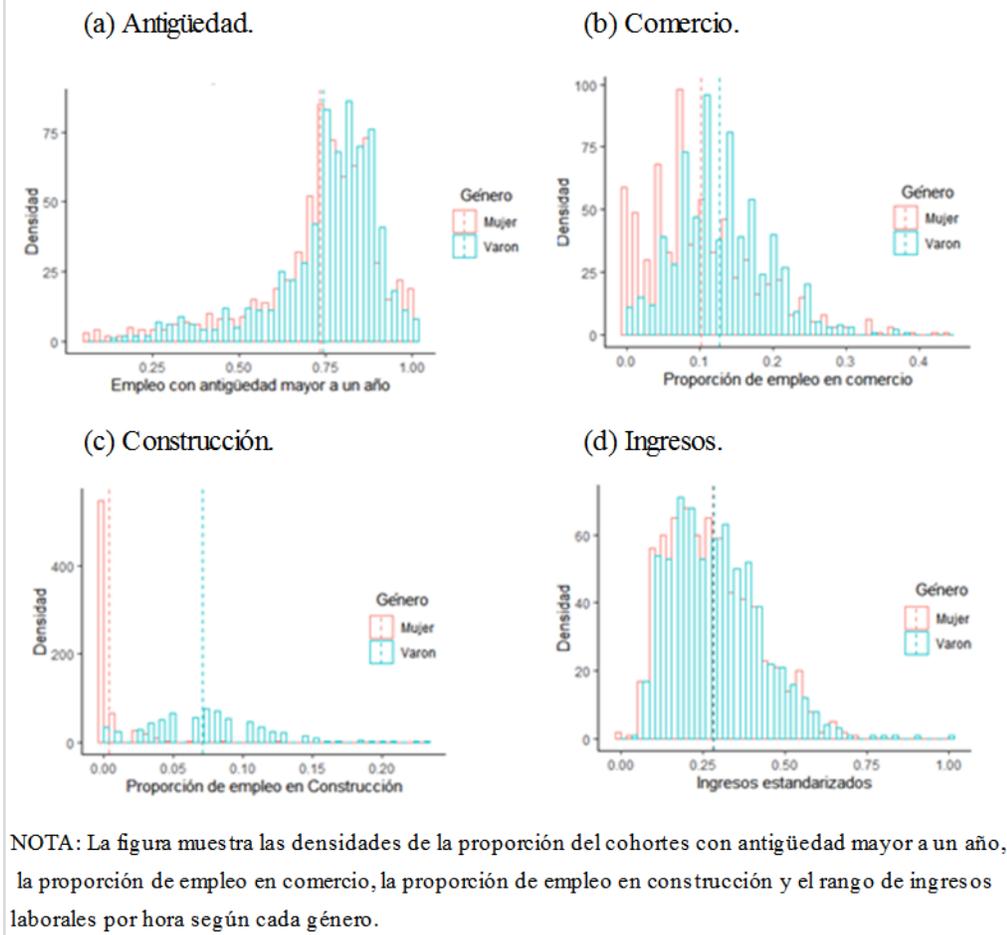
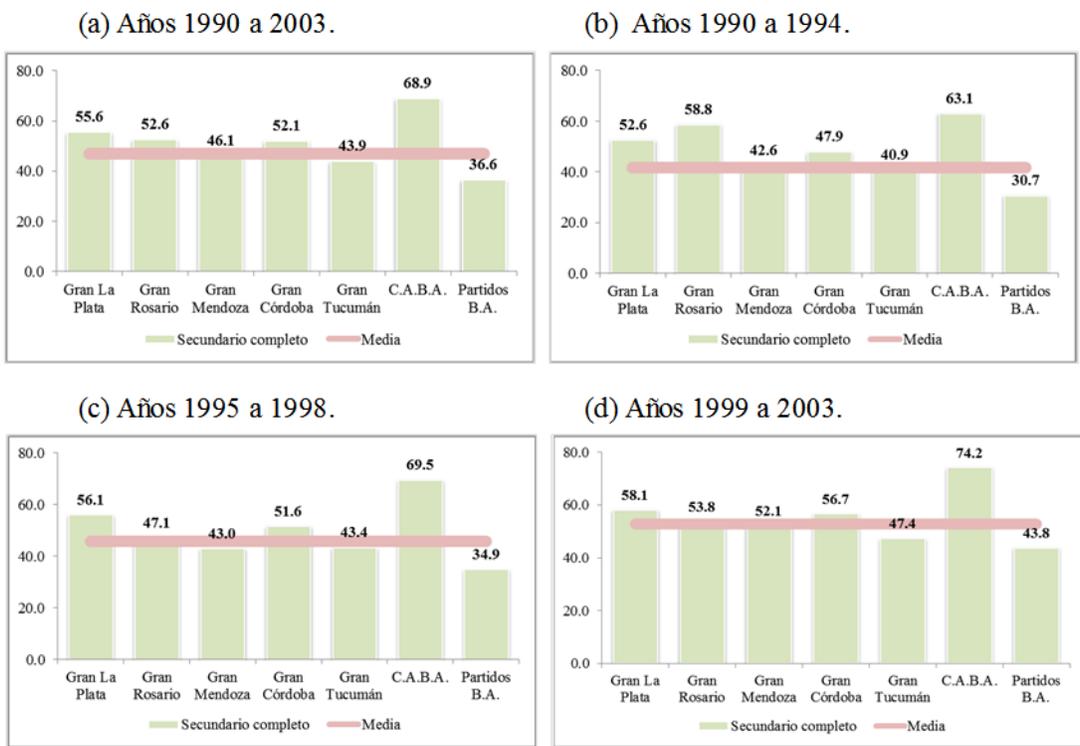
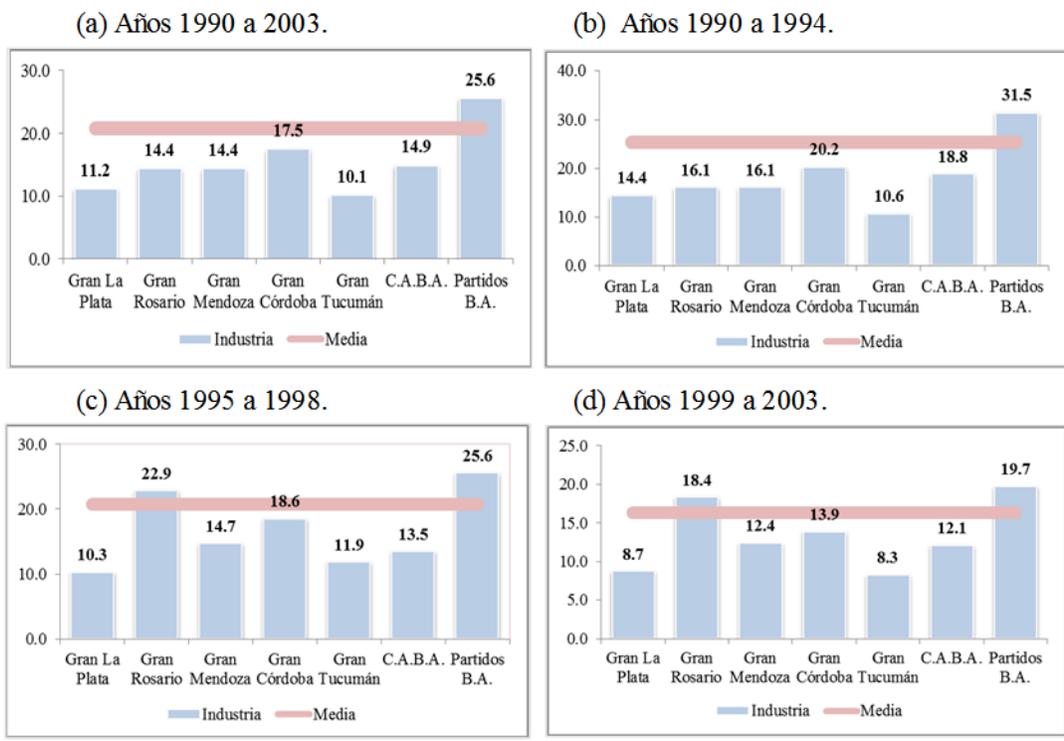


FIGURA B.3. Proporción de asalariados con educación secundaria completa.
Grandes centros urbanos. Argentina.



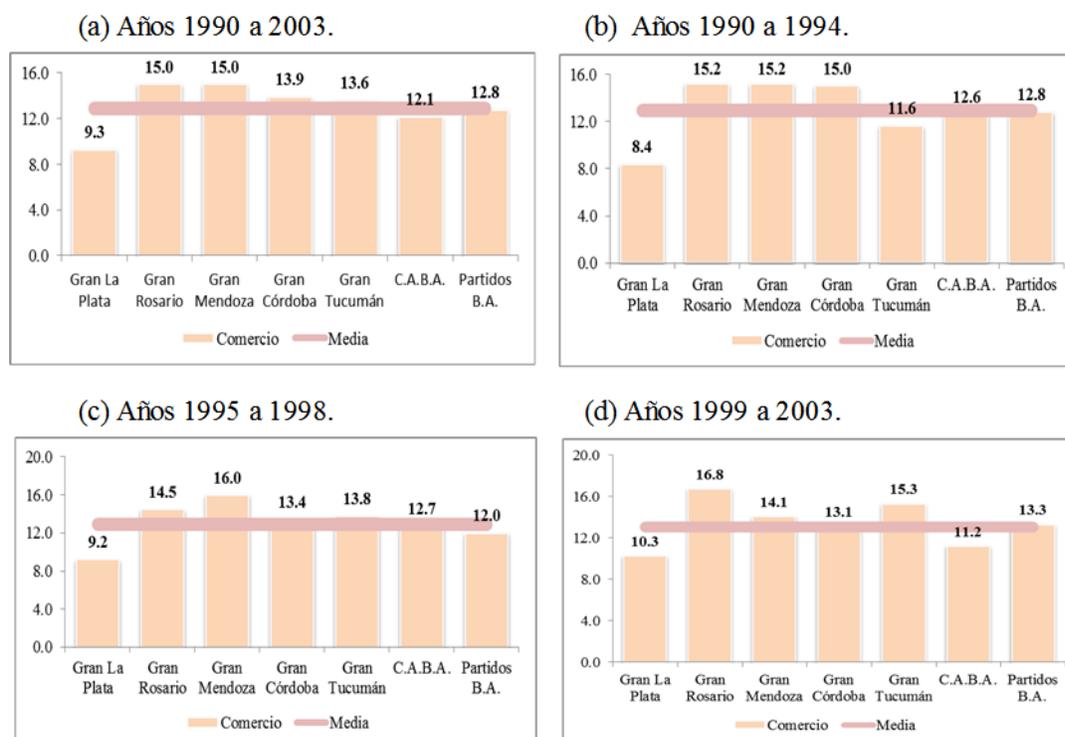
NOTA: Las barras indican la proporción de empleo asalariado con educación secundaria completa en cada aglomerado y la línea muestra la proporción promedio del conjunto de aglomerados en cada etapa.

FIGURA B.4. Proporción de asalariados con empleo en actividad industrial.
Grandes centros urbanos. Argentina.



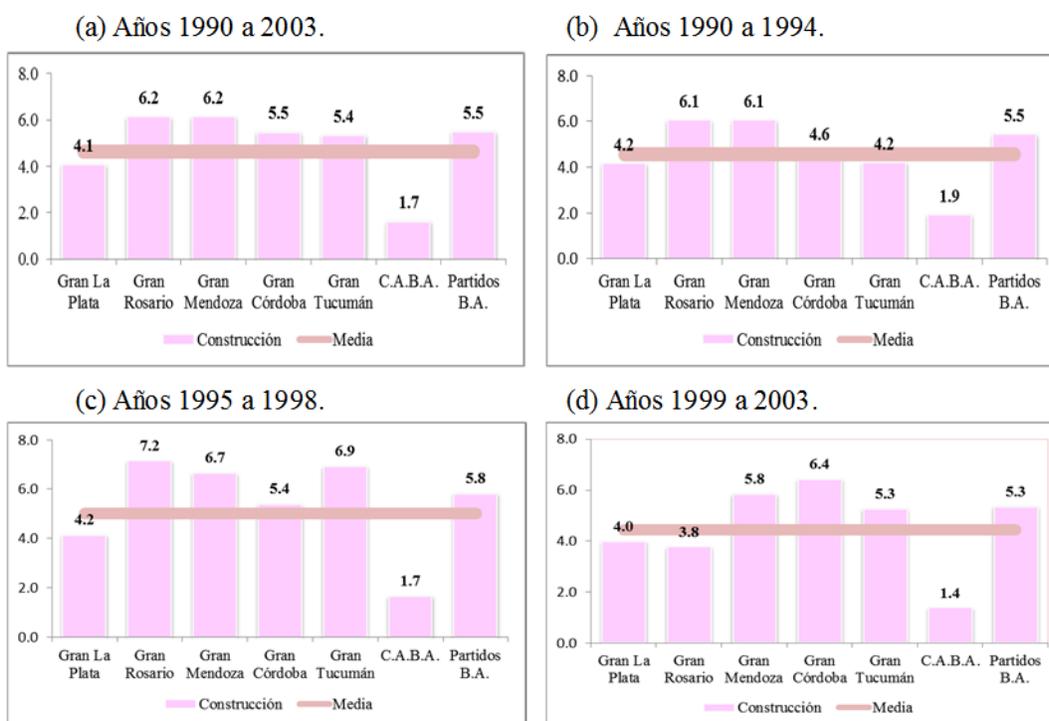
NOTA: Las barras indican la proporción de empleo asalariado en actividades de la rama industrial en cada aglomerado y la línea muestra la proporción promedio del conjunto de aglomerados en cada etapa.

FIGURA B.5. Proporción de asalariados con empleo en actividad comercial.
Grandes centros urbanos. Argentina.



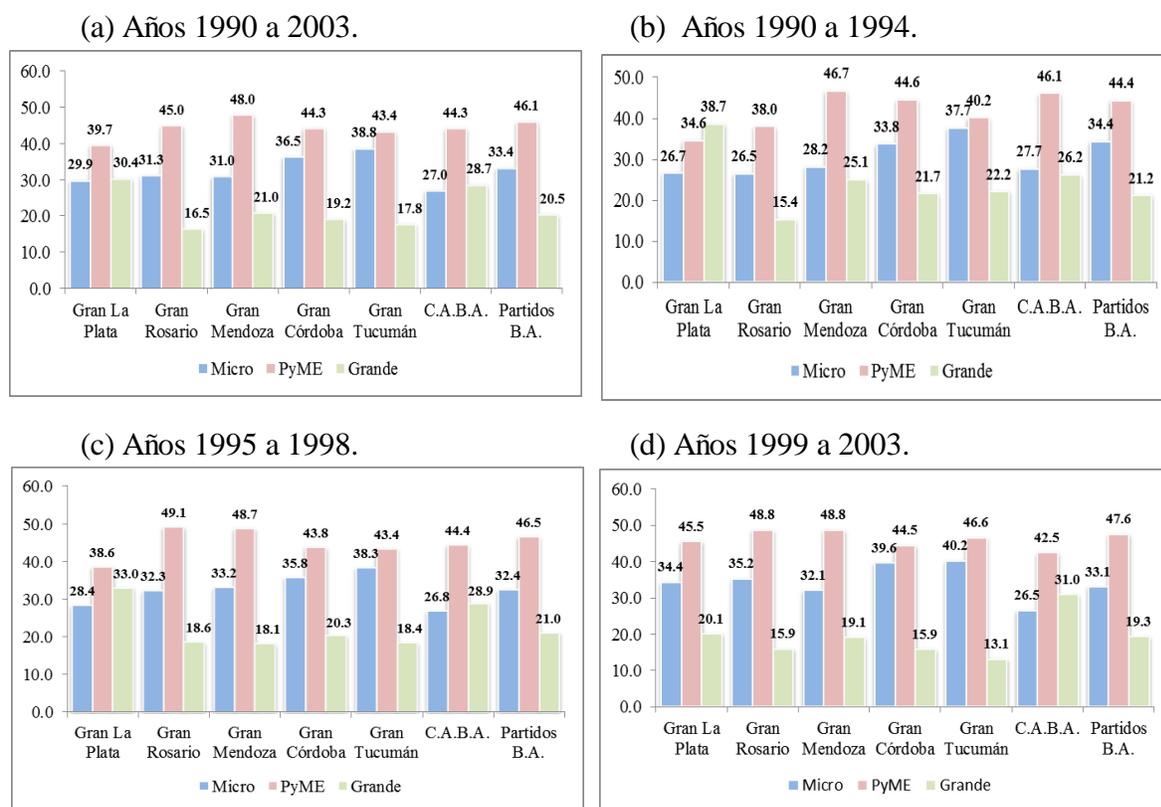
NOTA: Las barras indican la proporción de empleo asalariado en actividades de la rama de comercio en cada aglomerado y la línea muestra la proporción promedio del conjunto de aglomerados en cada etapa.

FIGURA B.6. Proporción de asalariados con empleo en actividad construcción.
Grandes centros urbanos. Argentina.



NOTA: Las barras indican la proporción de empleo asalariado en actividades de la rama construcción en cada aglomerado y la línea muestra la proporción promedio del conjunto de aglomerados en cada etapa.

FIGURA B.7. Proporción de asalariados según tamaño de establecimiento empleador.
Grandes centros urbanos. Argentina.



NOTA: Las barras indican la proporción de empleo asalariado según tamaño de establecimiento empleador en los aglomerados y la línea muestra la proporción promedio del conjunto de aglomerados en cada etapa.

Tabla B.9.

Medidas descriptivas. Variables en asalariados en el periodo 1990-2003.

VARIABLE	Media Aritmética	Desviación Estándar	Coefficiente Variación	Mínimo	Q1	Q2	Q3	Máximo	RI	Coefficiente Asimetría	Coefficiente Curtosis
Primario Incompleto	8.1	1.8	22.7	5.5	6.7	8.0	9.1	11.5	2.3	0.2	-0.6
Secundario Incompleto	53.1	5.3	9.9	44.5	49.0	53.8	56.9	59.3	7.9	-0.5	-1.1
Industria	20.8	4.1	19.5	15.2	17.0	20.6	24.6	26.9	7.6	0.1	-1.5
Comercio	12.9	0.7	5.1	11.8	12.5	12.7	13.4	13.9	0.9	-0.1	-0.9
Construcción	4.1	1.2	28.3	2.4	3.2	3.9	5.2	5.8	2.0	0.2	-1.5

Fuente: Elaboración propia en base a datos de la Encuesta Permanente de Hogares Puntual.

ANEXO IV: C. METODOLOGÍA.

Tabla C.1.

Cantidad de personas relevadas con edades entre 18 y 60 años en las encuestas transversales de grandes centros urbanos de Argentina según género. Muestra sin expandir.

AÑO	ONDA	VARONES	MUJERES	TOTAL
1990	1	16569	17849	34418
1991	2	16234	17092	33326
1992	1	17257	18702	35959
1993	1	17782	19123	36905
1994	1	17794	19222	37016
1995	1	17826	19197	37023
1996	1	17970	19136	37106
1997	1	17825	18894	36719
1998	1	17407	18343	35750
1999	1	15801	16965	32766
2000	1	15904	17154	33058
2001	1	15953	17024	32977
2002	1	15802	17079	32881
2003	1	10124	10765	20889

Tabla C.2.

Cantidad de asalariados relevados con edades entre 18 y 60 años en las encuestas transversales de grandes centros urbanos de Argentina según género. Muestra sin expandir.

ASALARIADOS

AÑO	ONDA	VARONES	MUJERES	TOTAL
1990	1	4928	2866	7794
1991	2	5718	3326	9044
1992	1	5626	3379	9005
1993	1	5558	3387	8945
1994	1	5559	3404	8963
1995	1	5430	3409	8839
1996	1	5386	3316	8702
1997	1	5528	3465	8993
1998	1	5542	3656	9198
1999	1	5053	3415	8468
2000	1	5005	3520	8525
2001	1	4866	3552	8418
2002	1	4482	3352	7834
2003	1	4620	3499	8119

ANEXO V: D. RESULTADOS.

Figura D.1.a

Trayectoria de la informalidad laboral en las cohortes según género.

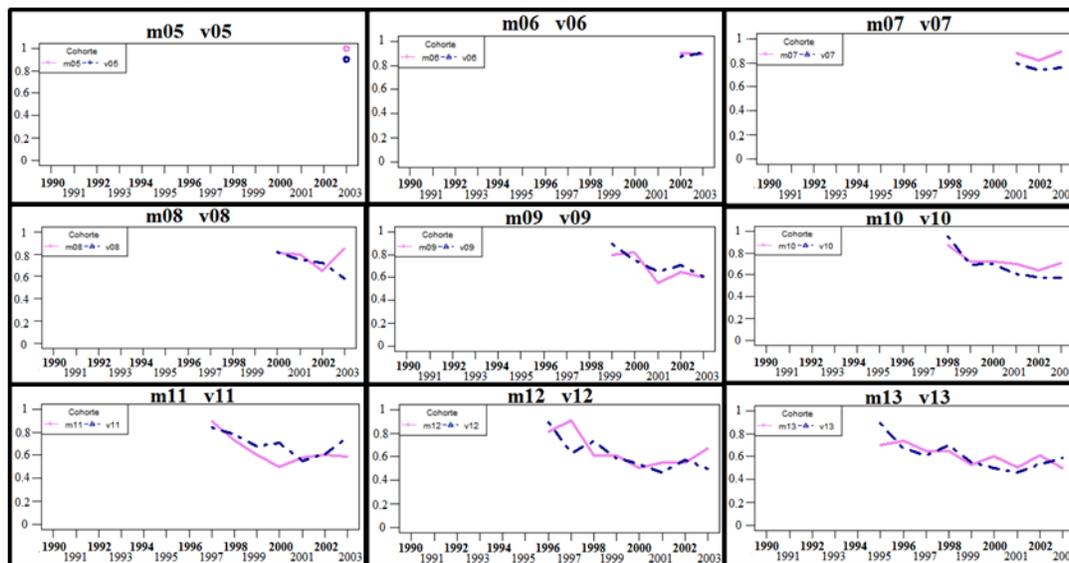


Figura D.1.b

Trayectoria de la informalidad laboral en las cohortes según género.

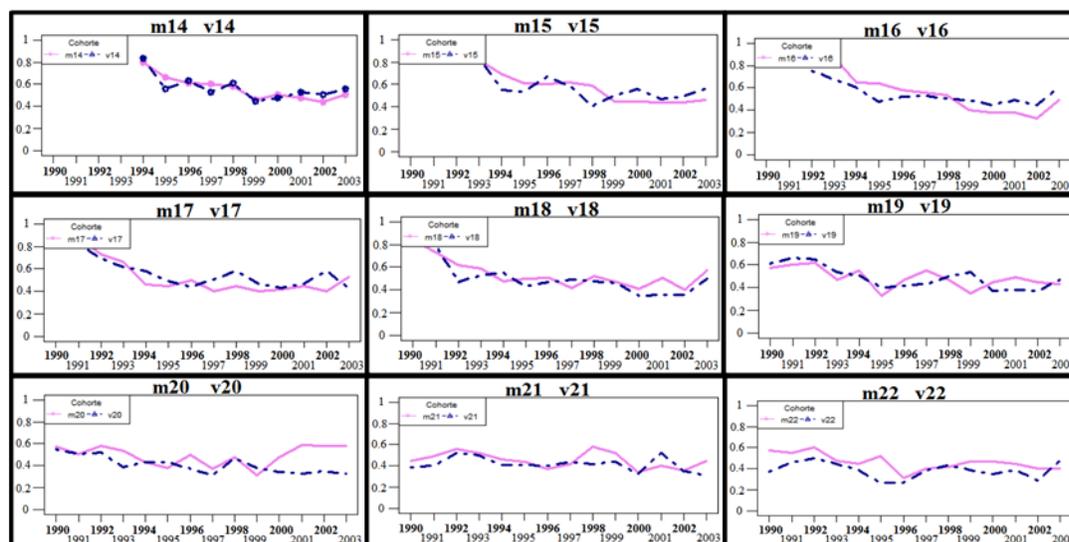


Figura D.1.c

Trayectoria de la informalidad laboral en las cohortes según género.

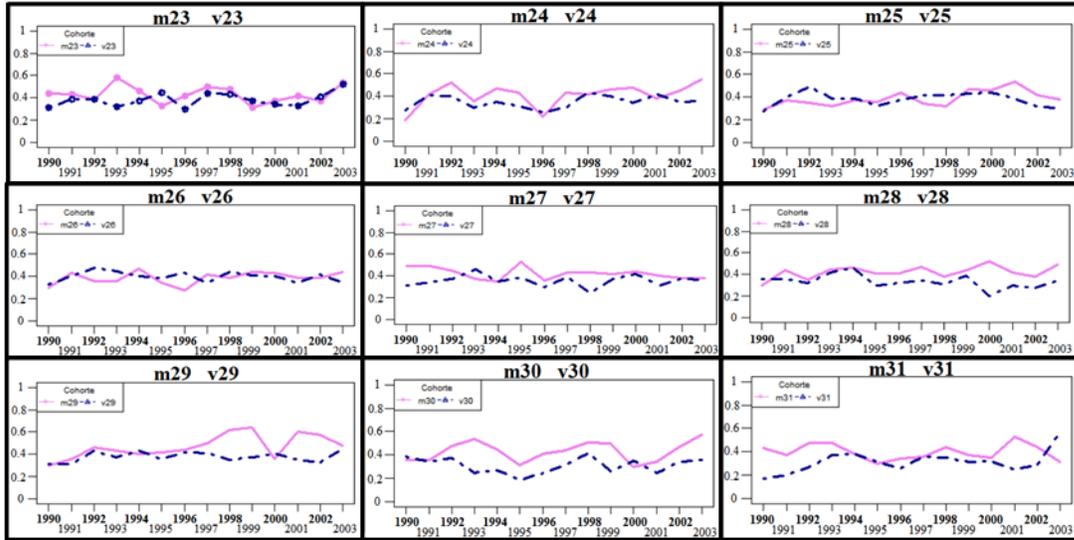


Figura D.1.d

Trayectoria de la informalidad laboral en las cohortes según género.

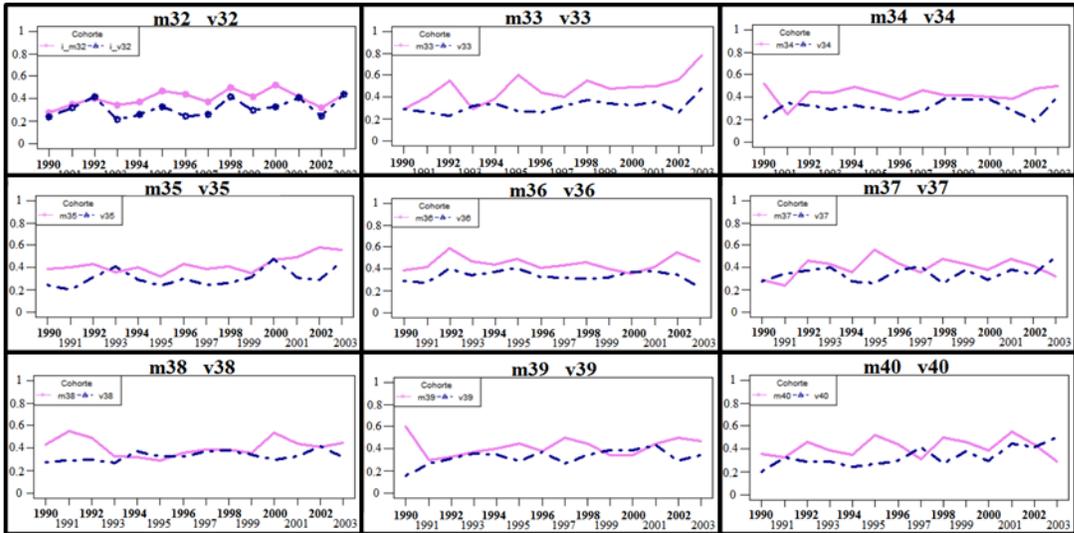


Figura D.1.f

Trayectoria de la informalidad laboral en las cohortes según género.

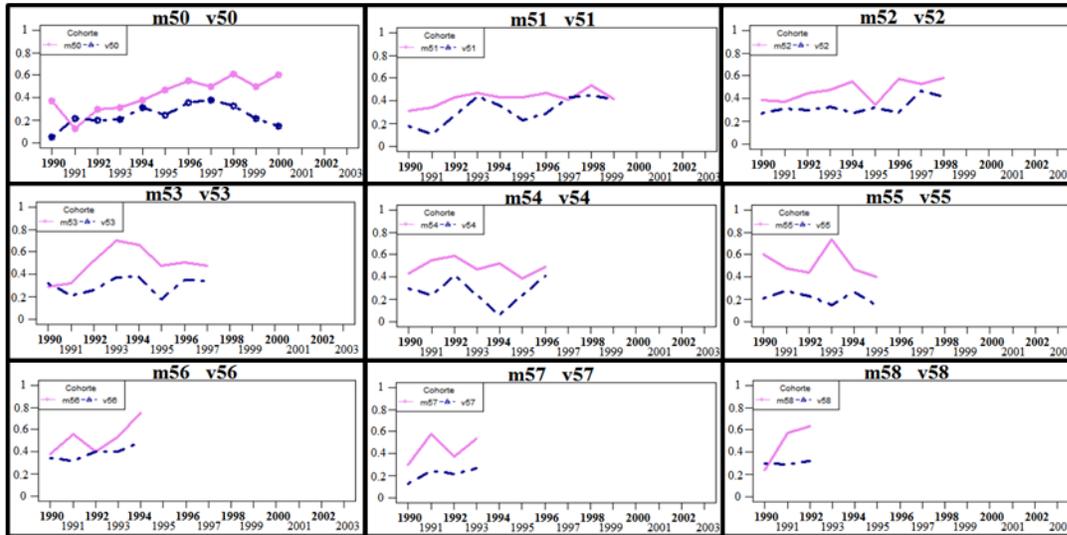
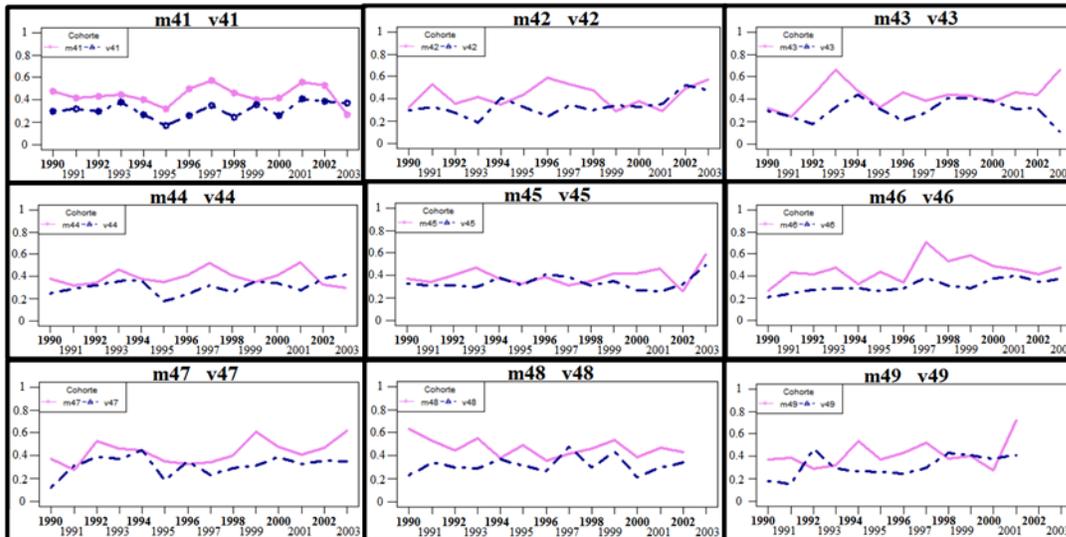


Figura D.1.e

Trayectoria de la informalidad laboral en las cohortes según género.



D.1. RESIDUOS DEL MODELO.

FIGURA D.2

Histograma de los residuos del modelo. Ajuste Normal (0,1).

Supuestos del modelo. Normalidad.

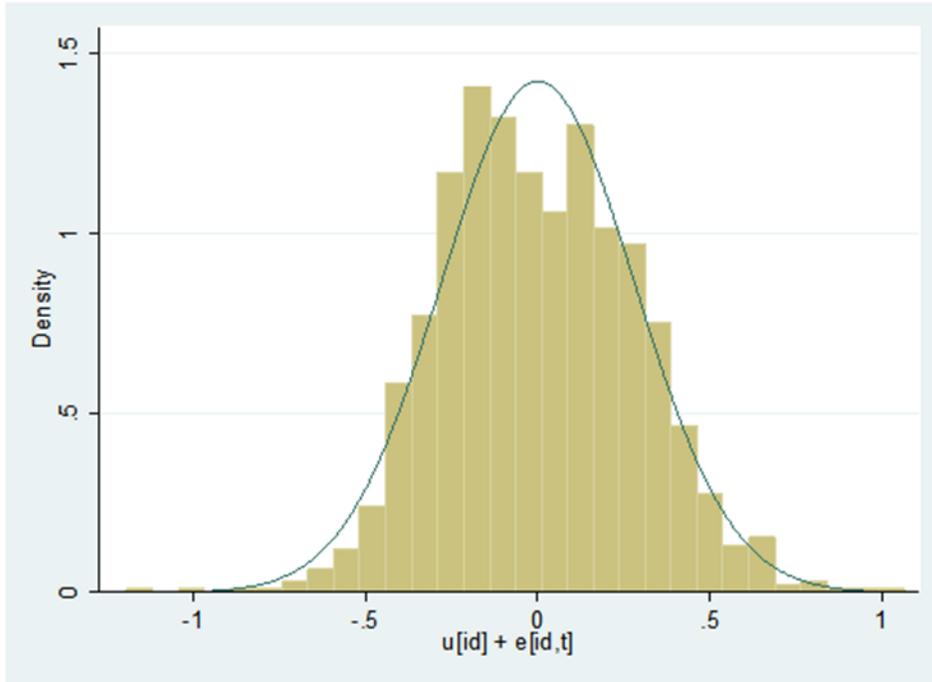


FIGURA D.3

QQ-Plot de los residuos del modelo. Ajuste Normal (0,1).

Supuestos del modelo. Normalidad.

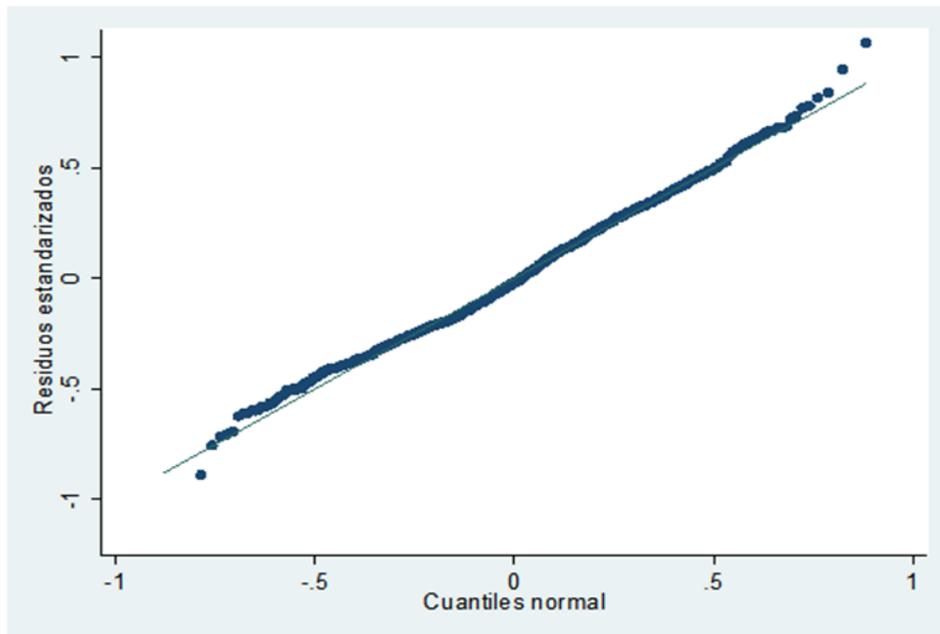


FIGURA D.4

*Gráfico de dispersión. Orden vs Residuos.
Supuestos del modelo. Independencia.*

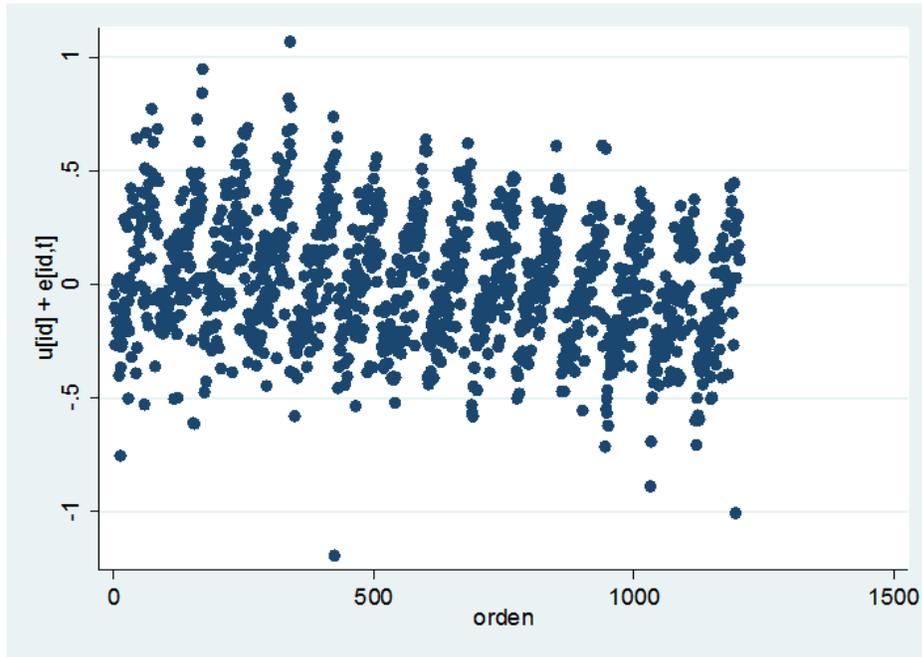
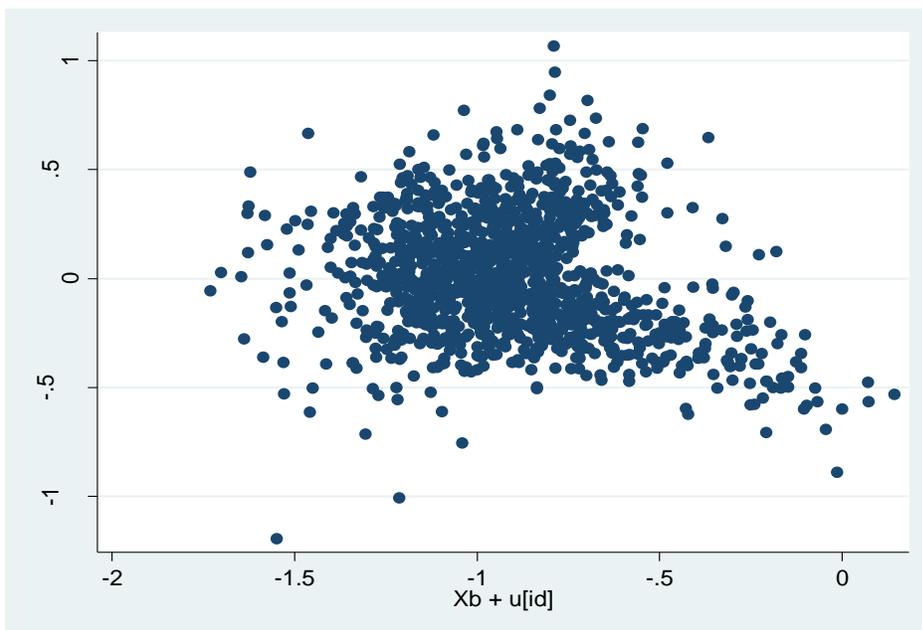


FIGURA D.5

*Gráfico de dispersión. Residuos vs Predichos.
Supuestos del modelo. Homocedasticidad.*



ANEXO VI: E. REPOSITORIO DE DATOS.

El siguiente enlace permite acceso a la carpeta “*REPOSITORIO_TESISMEAUNC_IGLESIAS*” en donde se encuentran contenidas las principales bases de datos y sintaxis utilizadas para realizar el presente trabajo.

<https://drive.google.com/drive/folders/1p3EAaiOXdUdqkBBMIewjtQbacGWPXzxL?usp=sharing>

La misma se encuentra dividido en tres sub-carpetas “*Bases de Datos*”, “*Scripts_Analisis*” y “*Scripts_Graficos*” y la descripción de los archivos alojado en cada una de ellas se indican en la Tabla E.1.

Posteriormente se pretende unificar los archivos e incorporar el trabajo en el formato de ciencia abierta y reproducible, R Script y R Markdown para una mejor lectura de los procedimientos utilizados.

Tabla E.1.

Descripción de archivos contenidos en el repositorio de datos.

CARPETA	NOMBRE ARCHIVO	DESCRIPCIÓN
Bases de Datos	BaseTrabajo.xlsx	<i>Base de datos de trabajo. Resultante de la construcción de cohortes. Previo a la modelación por enfoque de pseudo panel.</i>
	basepsp_media.xlsx	<i>Base de datos de valores medios de las variables "ingreso monetario por hora" y "niveles de informalidad" según año y cohorte definida.</i>
	basepsp_predichos.csv	<i>Base de datos de los predichos resultantes del modelo de pseudo-panel. Previo a la aplicación de la técnica K-mean Longitudinal.</i>
Scripts_Analisis	CodPaneles.1990a1994.sps	<i>Ejemplo sintaxis armado de cohortes. Válido para modelo de EPH Puntual utilizado entre 1990 y 1994.</i>
	A_Descriptivo.R	<i>Análisis descriptivo univariado (Parte III)</i>
	Test.R	<i>Pruebas univariados (Parte III)</i>
	PrepBase.do	<i>Selección de variables para modelo (Parte V)</i>
	ModSel.do	<i>Selección de modelos (Parte V)</i>
	Posest.do	<i>Análisis posestimación (Parte V)</i>
	KmL_media.R	<i>KmL aplicado sobre valores medios de pseudo-panel.</i>
	KmL_pedichos.R	<i>KmL aplicado sobre predichos de pseudo-panel.</i>
DescriptivoGrupos.sps	<i>Análisis descriptivo de cluster (Parte V)</i>	
Scripts_Graficos	Graficos_IIIa.R	<i>Gráficos (Parte III).</i>
	Graficos_IIIb.R	<i>Gráficos (Parte III).</i>
	Graficos_D.R	<i>Gráficos (Anexo D).</i>