

CAPÍTULO 2

Metodología para el estudio de la creación de empresas

Ana Gennero, Fernando Graña y Patricia Alegre

INTRODUCCIÓN

Las metodologías tradicionales para el análisis de la creación de empresas, utilizadas generalmente en los estudios de dinámica industrial, se basan en la comparación intercensal o impositiva de registros de empresas, y tienen como finalidad identificar los patrones de natalidad, mortalidad y supervivencia de los nuevos emprendimientos (Dunne *et al.*, 1988; Mac Donald, 1987; Graña, 2000). A partir de allí es posible ahondar en los factores que inciden en el inicio del nuevo negocio mediante entrevistas con los empresarios. Sin embargo, este análisis retrospectivo suele generar dificultades de distinta índole, tales como distorsiones en la información, provocadas por la falta de memoria o la *post-hoc* racionalización de los hechos ocurridos tiempo atrás en el proceso de gestación.

Una de las falencias de estas metodologías es que no permiten captar a las empresas que nacieron y murieron en el período intercensal bajo estudio, ni analizar los casos de ideas y esfuerzos realizados en proyectos que no llegaron a concretarse. Asimismo, se corre el riesgo de excluir del análisis a las nuevas empresas no registradas, que en esta etapa se presume son un número elevado.

Detrás de su aplicación subyace la idea de que las empresas se crean en el momento en que comienzan a operar en el mercado, sin considerar que el proceso de creación es mucho más amplio y complejo. De allí la falta de

interés por el estudio de los períodos previos al inicio de actividades, es decir, de las etapas de concepción y gestación de los proyectos empresariales.

En la presente investigación se utiliza una metodología nueva. Ella hace posible estudiar la creación de empresas como un proceso, que se inicia cuando se concibe una idea de negocios y que avanza hasta que la empresa creada se halla consolidada en el mercado. A su vez, facilita el estudio de dicho proceso en el momento en el que se está desarrollando, y, por lo tanto, resuelve los inconvenientes derivados del uso de la memoria y de la racionalización *ex-post* de los hechos ocurridos en el pasado.

En este capítulo se presentan, en primer lugar, los pasos del diseño metodológico que responden a las etapas delineadas conceptualmente para el estudio del proceso de creación. En segundo lugar, se dan a conocer las ciudades en las que se realizó el trabajo y los criterios de su selección. Los dos últimos apartados se destinan a la metodología del diseño muestral, la elaboración de los estimadores poblacionales para el grupo objetivo y la construcción del grupo control empleado en el estudio de las características de los emprendedores nacientes.

I. ETAPAS DEL DISEÑO METODOLÓGICO

Los lineamientos básicos de la metodología aplicada fueron desarrollados en el *Babson College*, EE.UU. (Reynolds & Miller, 1992; Reynolds *et al.*, 1993; Reynolds, 1997). Los miembros del ERC (*Entrepreneurial Research Consortium*) aplicaron esta metodología en forma experimental, posteriormente, fue utilizada en una comparación internacional en el proyecto conjunto del *Babson College* y *London Business School*, GEM (*Global Entrepreneurship Monitor*).

La misma responde a las bases teóricas presentadas en el capítulo anterior y hace posible seguir el proceso de creación de firmas a lo largo de sus distintas etapas evolutivas. Se parte de la detección en la población de aquellos adultos que están realizando acciones concretas para la implementación de un proyecto empresarial, a fin de constituir un panel de emprendedores en

gestación cuyo comportamiento y evolución serán seguidos durante un lapso de tres años.

A. Primera etapa

En esta fase se detectan dentro de la población de una ciudad, región o país, a las personas adultas que, teniendo motivación y una idea para iniciar un negocio, están avanzando en las etapas necesarias para su implementación, sin haber logrado aún dar nacimiento a una empresa. A tales individuos se los denomina Emprendedores Nacientes (EN).

Para llevar adelante la primera etapa metodológica, se realizó un muestreo poblacional en cinco ciudades argentinas. A partir de la aplicación de un cuestionario estructurado y de la realización de entrevistas personalizadas, se obtuvo información de los individuos adultos –mayores de 18 años–, representativos de la población de cada aglomerado, en un momento en el tiempo.

El objetivo es doble: en primer lugar, detectar a los EN y estimar la tasa de gestación empresarial de cada una de las ciudades bajo estudio. En segundo lugar, conocer las características socio-económicas de los emprendedores y de los hogares en los que habitan, para establecer qué aspectos los diferencian del resto de la población. Para ello se solicita información de dos tipos, la cual se describe en los apartados siguientes.

A.1. Detección de Emprendedores Nacientes

A cada uno de los miembros mayores de 18 años que habitan al interior de las viviendas seleccionadas en la muestra se les pregunta si ellos, solos o con otros, están intentando actualmente iniciar una nueva empresa:

- 1) Si la respuesta es afirmativa, se les consulta si han realizado dentro de los últimos doce meses actividades concretas para avanzar en la implementación de la idea.

- 2) Si responden afirmativamente a la pregunta anterior se les solicita un detalle de las actividades iniciadas o completadas hasta la fecha, en el proceso de gestación de la firma.

- 3) Si el individuo ha completado por lo menos dos actividades propias del proceso de gestación, que representen acciones concretas para lograr la implementación de la idea, se lo considera un EMPRENDEDOR NACIENTE y pasa a conformar el panel de la presente investigación¹.

El diseño muestral realizado para obtener información sobre los emprendedores nacientes permite estimar la TASA DE GESTACIÓN EMPRESARIAL en la población de cada una de las ciudades bajo estudio, con un error máximo de 10%.

A.2. Datos socio-económicos y de la vivienda

Una vez detectados los EN representativos de la población de cada ciudad bajo estudio, se recolecta información acerca de las características socio-ocupacionales de todos los individuos mayores de 18 años que habitan en los hogares seleccionados. Aquellos individuos que no reúnen las condiciones para ser considerados EN, son incluidos en un grupo control que representa a la población de cada ciudad que no tiene actualmente ideas empresariales en gestación.

A todos los integrantes de los hogares muestreados se les solicita la información respecto de:

- 1) Condición de actividad: ocupado, desocupado o inactivo.
- 2) Categoría ocupacional: asalariado, cuentapropista o empleador.
- 3) Estabilidad ocupacional: trabajo permanente o transitorio.
- 4) Nivel de educación formal.

¹ La empresa se considera iniciada o en operación cuando ha realizado ventas en forma regular o ha obtenido un *cash-flow* positivo por más de 3 meses. Esto es, los emprendedores que reúnan alguna de estas condiciones no son considerados EN sino empresas en funcionamiento.

Asimismo, se recaba la siguiente información respecto de la vivienda y de sus componentes:

- 1) Vivienda: tipo y propiedad.
- 2) Hogar: número de miembros y relación de parentesco.
- 3) Variables demográficas de los miembros: edad, género y estado civil.
- 4) Niveles de ingreso.

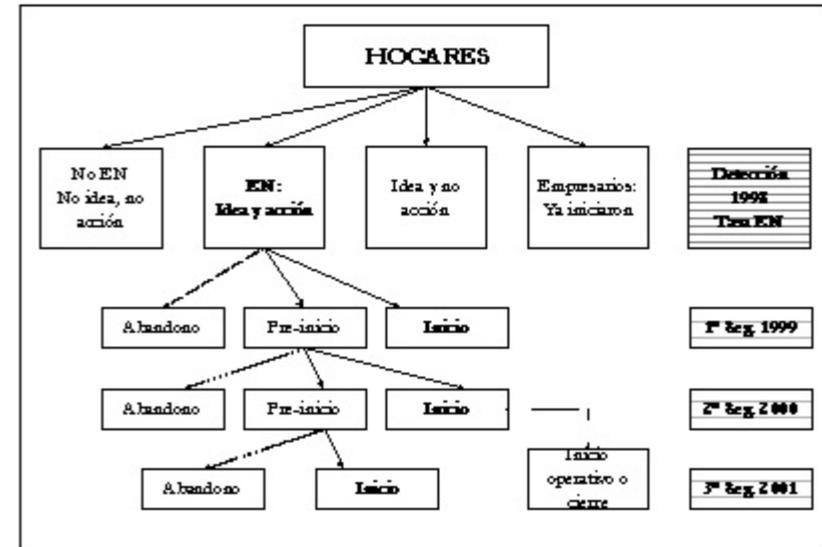
B. Segunda etapa

La segunda etapa metodológica consiste en seguimientos de los EN que conforman el panel, efectuados al primer, segundo y tercer año, los cuales brindan información sobre las distintas etapas del proceso de creación. Durante cada seguimiento se pueden identificar tres subgrupos:

- 1) Aquéllos que abandonaron la idea o no realizaron en ese período ninguna acción concreta para implementarla (abandonos).
- 2) Aquéllos que avanzaron con el emprendimiento pero todavía no han iniciado sus actividades (pre-inicios).
- 3) Aquéllos que iniciaron la empresa (inicios).

Las distintas fases de la metodología empleada se sintetizan en la Figura 2-1. A ella se hará referencia en los capítulos subsiguientes.

FIGURA 2-1
Metodología para la detección y seguimiento de EN



Fuente: Elaboración propia.

En el primer seguimiento (año 1999) se define al subgrupo (1) como grupo control y se lo compara con el (3), a fin de establecer los factores involucrados en el éxito o fracaso durante el pasaje de la etapa de gestación a la de infancia. En esta fase se estudian tanto las características diferenciales del emprendedor, sean éstas innatas o adquiridas en el medio socio-cultural de referencia, como las acciones realizadas para implementar el nuevo negocio, los tiempos insumidos y las inversiones realizadas, entre otros aspectos. Este análisis se presenta en los Capítulos 5 y 6.

A los 24 meses (noviembre 2000) se realiza un segundo seguimiento en el cual se estudia la evolución de aquellos emprendimientos que aún continúan en gestación. En esta instancia se focaliza el relevamiento sobre las empresas que se iniciaron en el año anterior, siendo el objetivo conocer la permanencia de las firmas en el mercado y su desempeño durante el primer año de vida. Los resultados se incluyen en el Capítulo 7.

Para analizar la permanencia, se clasifica a las empresas iniciadas en 1999 según continúen en actividad o hayan cesado durante el primer año de vida. Estas últimas son consideradas como grupo control para identificar los factores diferenciales relacionados a la supervivencia empresarial.

Para conocer el desempeño de aquellas empresas que permanecieron en actividad, se les solicita una evaluación de su *performance* en el primer período de funcionamiento, siendo las alternativas Bueno/Muy bueno y Regular/Malo. Sobre la base de las respuestas obtenidas se establecen dos grupos de empresas para analizar los factores asociados a dichos resultados.

Finalmente, entre los años 2001 y 2002, luego de tres años de la detección se realiza un último seguimiento que permite conocer la evolución general del panel de EN durante las distintas etapas del proceso de creación de una nueva firma. En esta última instancia se centra la atención sobre quienes no han podido implementar su proyecto y sobre quienes han podido sobrevivir a este período inicial tan traumático para la mayoría de las nuevas firmas, particularmente si el contexto macroeconómico es sumamente adverso. Los resultados se presentan en el Capítulo 8.

II. ALCANCE GEOGRÁFICO DEL ESTUDIO

La elección de las ciudades a estudiar –Córdoba, Mar del Plata, Bahía Blanca, Tandil y Rafaela– se efectúa atendiendo a la combinación de distintos factores. Los aspectos contemplados en la selección son la diversidad de tamaño, de estructuras productivas y el grado de dinamismo institucional, así como la cultura *entrepreneurial* de cada ciudad.

La ciudad de Córdoba es la segunda ciudad más grande del país –1.300.000 habitantes–, tiene una importante participación de las actividades industriales en su estructura productiva con especialización en la producción de maquinarias, equipos y automotores. Durante los años previos al estudio se consolida el conglomerado automotor con la radicación de empresas transnacionales como *Fiat*, *Renault*, *General Motors*, *Chrysler*, *Volkswagen* (Transax) y la fábrica de camiones *Fiat Iveco*. Existe, además, un grupo de aproximada-

mente 200 empresas autopartistas que abastecen a las terminales y al mercado de reposición nacional e internacional. Éstas elaboran productos tales como diferenciales, cajas de velocidad y amortiguadores, generando la instalación de industrias conexas de fabricación de maquinarias, herramientas, matrices y dispositivos diversos (Bolsa de Comercio de Córdoba, 1998)².

Aún cuando estudios locales hablan de la escasa formación de redes empresariales y de la falta de incentivos a la capacitación, de entrenamiento de RR.HH. y de una estrategia regional de desarrollo, la ciudad cuenta con un Plan Estratégico (PEC) y una Agencia para el Desarrollo Económico (ADEC) creada en 1997 con la participación de organismos públicos y privados y con objetivos claros de promoción de las actividades económicas.

Las actividades se agrupan en cuatro proyectos fundamentales: (a) apoyo a pequeños y medianos empresarios que incluye financiamiento y capacitación empresarial y laboral; (b) promoción internacional y apoyo al comercio exterior; (c) información para el desarrollo con un observatorio económico-ocupacional y (d) desarrollo territorial con un polígono industrial y un tecnoparque. Al interior de cada programa hay acciones específicas para la promoción del desarrollo local (Argañaraz, Olivero y Garzón, 1999).

Las ciudades de Mar del Plata y Bahía Blanca se seleccionaron por su tamaño mediano –550.000 y 280.000 habitantes respectivamente–, presentando ambas recientes esfuerzos de apoyo a las actividades productivas por parte de las instituciones locales. Mar del Plata tiene una estructura productiva basada en el sector servicios, especialmente ligados al turismo, esperándose, por lo tanto, una alta capacidad de generar nuevos emprendimientos. Ello es así por la evolución favorable de los precios relativos servicios/industria, dado en general más por el poder de mercado (ligado a la baja apertura a la competencia externa, escasa movilidad geográfica) que por el crecimiento de la

² A partir de 1998, el desplome de la producción automotriz perjudicó a la industria autopartista. En la Ciudad de Córdoba, dos terminales automotrices –*General Motors* y *Chrysler*– levantaron sus plantas y *Fiat* suspendió la producción de vehículos, situación que perdura en el 2002, en el que se finaliza el estudio (Motta y Zavaleta, 2004).

productividad, haciendo más atractivos dichos emprendimientos. A su vez, el hecho de tener el sector servicios mayores requerimientos de mano de obra que de capital, disminuye las barreras de entrada y facilita el ingreso al mismo (García Reche, 2003).

A su vez, Bahía Blanca posee una muy elevada participación del sector industrial en su producto bruto (43,7%), con alta especialización en la rama petroquímica y niveles de concentración del producto en las cuatro empresas de mayor tamaño que ascienden al 80%. Estas características de su estructura productiva hacen esperar una baja tasa de creación de nuevas firmas, dados los mayores salarios que suelen pagar las grandes empresas, así como la falta de nexos de subcontratación y la menor posibilidad de desarrollar capacidades empresariales por parte de los empleados en las firmas de esta rama industrial (Mason, 1991; Fothergill & Gudgin, 1982; Glasmeier, 1988).

Por último, se seleccionaron como ciudades de tamaño pequeño a Tandil y Rafaela –menos de 100.000 habitantes–. Ambas tienen una elevada participación del sector agropecuario en su estructura productiva –29,5% y 21,9%, respectivamente–, especialmente relacionada a la producción láctea.

Rafaela tiene una alta participación de las actividades industriales en el producto bruto local –40,1% vs. 31,1% en Tandil–. Se manifiesta allí una fuerte especialización en las ramas alimenticia –láctea y frigorífica– y metalmeccánica orientada a la producción de maquinaria para el agro y para la industria alimenticia, concentrando ambas ramas de actividad el 45% del producto bruto manufacturero y terciario local.

Las empresas alimenticias son de gran tamaño y muy pujantes; la producción de la industria metalmeccánica proviene predominantemente de pequeñas y medianas firmas, muchas de las cuales comenzaron como prestadoras de servicios de reparación y mantenimiento (Plan Estratégico de Rafaela, 1997). Sin embargo, los niveles de subcontratación en la industria local son bajos, estimándose que el 21% de las empresas destinan más del 25% de su producción a terceros (Municipalidad de Rafaela, 1997).

Por su parte, Tandil está especializada en la producción metalmeccánica –matricería, cilindros, radiadores y partes de automotores– la cual ha sufrido

una disminución de producción en las últimas décadas, representando en 1994 el 22% del producto bruto manufacturero y terciario. No obstante, Tandil tiene una universidad especializada en temas agrarios e ingeniería de sistemas, de las cuales se espera puedan surgir emprendimientos relacionados tanto a la tecnología de sistemas como a la biotecnología o a la producción agroalimentaria. A su vez, tiene una incipiente pero creciente actividad ligada al turismo.

La estructura institucional de esta ciudad se encuentra menos desarrollada que la de Rafaela. Esta última ha sido en los últimos 20 años un caso testigo en cuanto a dinamismo institucional y proyectos público-privados de inserción internacional –décadas del 80 y 90– que le permitieron a la ciudad mantener el nivel de actividad económica aún en períodos de crisis. Además, la formulación e implementación del primer Plan Estratégico en el año 1996, el cual genera un proyecto integral de desarrollo local.

En Tandil, recién en la última década se han realizado esfuerzos con miras a constituir un entramado institucional de apoyo al desarrollo del sector tecnológico-productivo. La selección de estas ciudades intenta, de esta forma, contemplar el impacto que tienen entornos institucionales de diferente grado de desarrollo sobre las actividades de creación de nuevos proyectos empresariales.

III. DISEÑO MUESTRAL PARA DETECCIÓN DE EMPRENDEDORES NACIENTES Y ESTIMADORES POBLACIONALES

En este apartado se presentan los pasos seguidos en la selección de los hogares a ser entrevistados para la detección de Emprendedores Nacientes. La metodología utilizada difiere entre las ciudades de Córdoba, Mar del Plata y Bahía Blanca, las cuales pertenecen a la muestra nacional de la Encuesta Permanente de Hogares (EPH) y cuentan con un marco muestral de

actualización permanente, y las ciudades de Rafaela y Tandil no incluidas en la EPH³.

En las ciudades incluidas en la EPH se ha utilizado la muestra existente para un primer relevamiento, realizándose luego una ampliación muestral. Esta necesidad de inclusión de un mayor número de hogares en la muestra se debe a la baja tasa de EN estimada en la prueba piloto que, con el tamaño muestral inicial, impide tener un número significativo de casos a ser analizados, incrementando el coeficiente de variación por encima del 10%.

En las ciudades de Rafaela y Tandil, al no existir al momento del relevamiento de detección un marco muestral ni una muestra utilizada en relevamientos poblacionales de carácter permanente, se diseñó una muestra específica⁴. Ello se explica en la Sección B.

El tamaño muestral por ciudad en el presente estudio oscila entre 2.100 y 2.400 hogares. La metodología de muestreo consiste en una muestra probabilística bietápica, tomada a partir de los hogares de las ciudades seleccionadas en las que el INDEC realiza la EPH y un muestreo simple al azar para las restantes ciudades.

A. Aglomerados con marco muestral de la EPH

En la onda de encuestas de agosto de 1998 realizada por el INDEC en los aglomerados de Bahía Blanca, Córdoba y Mar del Plata, se anexó al formulario de la EPH un conjunto de preguntas, con el objeto de identificar a los EN de esas localidades. Posteriormente, en el mes de noviembre, se realizó un relevamiento adicional para completar la muestra de detección de EN, ampliando la muestra a partir del mismo marco muestral pero contemplando un mayor número de viviendas. A continuación se presenta la

³ La EPH es realizada por el INDEC (Instituto Nacional de Estadísticas y Censos).

⁴ En este diseño se contó con la colaboración de técnicos del Departamento de Muestreo del INDEC (octubre, 1998).

metodología utilizada por el INDEC para el diseño de la muestra de hogares a ser encuestados dentro de la EPH y la metodología usada para la ampliación de la misma.

A.1. Determinación del marco muestral: EPH

La incorporación de un aglomerado a la EPH parte de los datos provistos por los censos de población, que divide a cada localidad en fracciones y radios censales. El censo es tomado como base para la organización del marco muestral de la encuesta y se considera que un radio tiene un tamaño promedio de 300 viviendas. Cada radio censal proporciona información básica sobre cantidad de viviendas, de hogares y del total de población clasificada por género⁵. Para una mayor precisión se actualizan las cifras censales realizando recorridos en terreno y delimitando así la zona o envolvente del aglomerado.

Un aglomerado se define como un conjunto de radios censales que en ocasiones incluye más de una localidad, no necesariamente contiguas geográficamente, pero relacionadas de manera tal que constituyen una sola unidad socio-económica (e.g., Mar del Plata - Batán). El criterio de selección del marco muestral difiere según el tamaño del aglomerado: si es inferior a 50.000 habitantes se trabaja con un muestreo probabilístico en una sola etapa de selección, mientras que en los aglomerados con una población mayor, el muestreo se realiza en dos etapas.

Previo a la selección de hogares, los marcos muestrales se estratifican con el fin de satisfacer las necesidades de información y reducir la variabilidad de las estimaciones. Las variables utilizadas en la estratificación son: porcentaje de jefes de hogar sin instrucción o con primaria incompleta y porcentaje de jefes de hogar con estudios universitarios o superiores completos –e.g., los valores

⁵ Se define vivienda como cualquier albergue, fijo o móvil, que pueda ser utilizado para alojar personas. Se define hogar como una persona o grupo de personas, emparentadas o no, que habitan bajo un mismo techo y se asocian para proveer a sus necesidades alimenticias u otras esenciales para vivir.

establecidos para el estrato 1 de la ciudad de Mar del Plata son <10% y >15%, respectivamente–.

Debido a que los aglomerados bajo estudio presentan una población mayor a los 50.000 habitantes, se trabajó con un muestreo probabilístico en dos etapas de selección. En la primera etapa, los radios censales se ordenan según una serpentina geográfica, o Unidades de Primera Etapa (UPE), y se selecciona la muestra con una probabilidad proporcional a su tamaño medido en cantidad de viviendas ocupadas.

Una vez seleccionadas las UPE se realiza un listado completo de viviendas en campo y sobre el mismo se selecciona en forma independiente dentro de cada UPE una muestra de *n* viviendas (unidades de muestreo). Esto se efectúa mediante un procedimiento de selección sistemática con arranque aleatorio, de forma tal de obtener una muestra autoponderada, es decir cada vivienda tiene igual probabilidad de selección. De esta manera quedan conformadas muestras de 881 viviendas para Mar del Plata, 1.167 para Córdoba y 799 para Bahía Blanca.

A.2. Selección del segundo anillo muestral: ampliación de la muestra

El objetivo de este diseño muestral es obtener una muestra más amplia que la utilizada por la EPH, debido al escaso número de Emprendedores Nacientes detectados en el relevamiento de octubre de 1997 en Mar del Plata (prueba piloto). La ampliación de la muestra se realiza para las ciudades de Bahía Blanca, Córdoba y Mar del Plata. El nuevo anillo muestral fue seleccionado sobre el marco del INDEC que habitualmente se utiliza en encuestas poblacionales.

CUADRO 2-1
Número de viviendas seleccionadas

Aglomerado	Primera detección Agosto 1998 EPH	Segunda detección Noviembre 1998 Muestra ampliada	Tamaño muestral total
Córdoba	1.167	1.364	2.531
Mar del Plata	881	1.624	2.505
Bahía Blanca	799	1.293*	2.092

* Corresponde a agosto de 1999.

Fuente: Elaboración propia.

A.3. Estimaciones

Dado que la falta de respuesta puede generar distorsiones, se introduce un corrector de “No Respuesta” dentro de las fórmulas utilizadas para las estimaciones. Su construcción depende de las causas que la originaron y su naturaleza permite la definición de dos grupos.

CUADRO 2-2
Causas de no respuesta

GRUPO 1	GRUPO 2
Ausencia momentánea	Casa de fin de semana
Ausencia temporal	Casa deshabitada o demolida
Rechazo total	Casa o departamento en construcción
Rechazo parcial	Vivienda establecimiento
Otras causas	

Fuente: Elaboración propia.

Las causas del Grupo 1 indican una no respuesta real mientras que las del Grupo 2 se deben a desajustes en el marco muestral. Las viviendas que responden sumadas a los hogares que no responden por las causas enumeradas en el Grupo 1, constituyen el conjunto de viviendas efectivas. El factor de corrección, calculado a nivel de UPE y estrato, se define como:

$$C_{ki} = \frac{\text{viviendas efectivas}}{\text{viviendas que responden}}$$

siendo $h =$ estratos, $i =$ UPE. La expansión de la muestra se realiza utilizando un factor diferente para cada estrato socio-económico, el que se calcula como:

$$\text{Factor de expansión} = \frac{\text{viviendas en la población}}{\text{viviendas muestrales}}$$

Las fórmulas para estimar los totales y las varianzas se definen de la siguiente manera:

Valor de la variable	x_{kij}
Viviendas que responden	n'_{ki}
Corrector de no respuesta	C_{ki}
Cantidad de estratos	L
Total de la UPE	m_k
Factor de expansión	F_k
Total estimado	$\sum_{k=1}^L F_k \sum_{i=1}^{m_k} C_{ki} \sum_{j=1}^{n_{ki}} x_{kij}$
Varianza estimada	$\left(\sum_{k=1}^L F_k \right)^2 \frac{m_k}{m_k - 1} \sum_{i=1}^{m_k} (x_{ki} - \bar{x}_k)$
h = estratos = 1, ..., L; i = UPE = 1, ..., m _h ; j = vivienda = 1, ..., n' _{hi}	

B. Aglomerados sin marco muestral

B.1. Diseño de la muestra

Tanto para la ciudad de Rafaela como para Tandil, el objetivo de los diseños muestrales fue efectuar un único relevamiento. Se emplean como

punto de partida padrones de viviendas que contienen o bien la totalidad o bien un porcentaje elevado de los domicilios particulares de ambas ciudades.

Con el propósito de estimar la proporción de EN en la población con un coeficiente de variación de alrededor del 10%, se obtiene una muestra de 2.400 viviendas particulares para la ciudad de Rafaela y de 2.100 para la ciudad de Tandil. Para determinar el tamaño de cada muestra se define:

Proporción de EN (media)	p
Coefficiente de variación	$CV = \frac{DS}{p}$
Desvío estándar	$DS = \frac{\sqrt{p(1-p)}}{n}$
Promedio de personas por vivienda	\bar{x}
Total de viviendas en la muestra	$n_v = \frac{n}{\bar{x}}$

Los valores resultantes se presentan en el Cuadro 2-3, pudiendo observarse en la última fila el número de viviendas incluidas en la muestra.

CUADRO 2-3

Datos muestrales correspondientes a Rafaela y Tandil

		RAFAELA	TANDIL
p	Proporción de EN	0	0
CV	Coefficiente de variación	0	0
\bar{x}	Promedio de personas por vivienda	3,00	3,42
n_v	Viviendas seleccionadas	2.398	2.100

Fuente: Elaboración propia sobre datos del Censo de Población 1991.

En la ciudad de Rafaela, debido a que el marco muestral utilizado presentaba alrededor de un 10% de direcciones incorrectas –por ejemplo, direcciones repetidas, esquinas, direcciones incompletas– y teniendo en cuenta la no respuesta, se elevó el tamaño de la muestra un 15%, alcanzando a 2.400 viviendas. Para garantizar la representatividad de la muestra en el caso de presentarse inconvenientes durante el relevamiento, la misma se conformó por tres submuestras independientes denominadas réplicas.

Si bien cada una de las réplicas es representativa de la totalidad de las viviendas, es necesaria su agregación para lograr la precisión propuesta. Cada una de ellas se selecciona aleatoriamente en forma sistemática y se encuentra constituida por 800 viviendas para el caso de Rafaela y por 700 viviendas para el caso de Tandil. Designando con:

N	Cantidad de viviendas en el marco
n	Tamaño de la muestra en viviendas
n_r	Cantidad de réplicas
K_r	Intervalo de selección de cada réplica

se obtiene un intervalo de selección para cada réplica igual a:

$K_r = \frac{N}{\frac{n}{n_r}}$	
Rafaela:	$19.173 / (2.400 / 3) = 19.173 / 800 = 23,966$
Tandil:	$31.921 / (2.100 / 3) = 31.921 / 700 = 45,601$

B.2. Corrección y expansión de los valores muestrales

Para expandir los valores muestrales se utiliza un factor de expansión simple definido como:

$F = \frac{K_r}{n_r}$	
Rafaela:	$23,966 / 3 = 7,989$
Tandil:	$45,601 / 3 = 15,200$

A dicho factor de expansión se le incorpora una corrección por no respuesta y un ajuste por demografía. Como ya fue mencionado para el caso de las ciudades de Bahía Blanca, Córdoba y Mar del Plata, las causas de no respuesta son divididas en dos grupos, lo que permite definir cuáles de estas viviendas no encuestadas forman parte de las viviendas efectivas junto a las viviendas que respondieron.

		RAFAELA	TANDIL
V_r	Viviendas que respondieron	1.587	1.668
V_e	Viviendas efectivas	328	1.836
C	Factor de corrección por no respuesta	1,207	1,101

Utilizando el mismo procedimiento se calcula, además, un factor de corrección por no respuesta C_c para la muestra de control:

	Grupo Control	RAFAELA	TANDIL
V_{rc}	Viviendas que respondieron	497	263
V_{ec}	Viviendas efectivas	644	302
C_c	Factor de corrección por no respuesta	1,296	1,148

Para mejorar las estimaciones, se realiza un ajuste por demografía utilizando una fuente de información externa –proyección demográfica (P_{ry}) por ciudad–. Sea Z_c el valor estimado del total de la población que resulta de

multiplicar el total de personas en la muestra (x) por el factor de expansión y por la corrección por no respuesta:

$$Z_e = C \cdot F \cdot x$$

el factor de ajuste por demografía (D) resulta: $D = \frac{P_{99}}{Z_e}$

Los factores de ajuste por demografía obtenidos fueron los siguientes:

		RAFAELA	TANDIL
x	Total de personas de 18 años y más en la muestra	3.520	385
Z_e	Total estimado de personas de 18 años y más	33.933	63.661
P_{99}	Proyección de la población de 18 años y más (año 1998)	54.844	66.959
D	Ajuste por demografía	1,616	1,000

B.3. Estimaciones

Para obtener una estimación del total (X_e) para una determinada característica de la población a partir del valor muestral x , se procede de la siguiente manera:

$$X_e = D \cdot C \cdot F \cdot x$$

La ponderación para el total de la muestra (W) y para el grupo control (W_c), respectivamente, se definen como:

$$W = D \cdot C \cdot F$$

$$W_c = D \cdot C_c \cdot F_c$$

Aplicando las fórmulas precedentes, los ponderadores utilizados fueron:

		RAFAELA	TANDIL
W	Ponderador para el total de la muestra	15,581	17,598
W_c	Ponderador para el grupo control	50,040	121,000

IV. EMPRENDEDORES NACIENTES Y GRUPO CONTROL

Se considera como población de control a aquellas personas de 18 años y más que no eran emprendedores nacientes al momento del relevamiento. Con la finalidad de constituir un GC, en la ciudad de Rafaela se selecciona al azar una de las tres réplicas muestrales y en la ciudad de Tandil se selecciona una submuestra de 350 viviendas dentro de la primera réplica. En Mar del Plata, Bahía Blanca y Córdoba, el GC queda constituido por las personas de 18 años y más que respondieron la EPH en la salida de agosto de 1998 y que manifestaron no ser emprendedores nacientes.

Estas muestras son revisadas en un paso posterior, para ajustar las diferencias existentes en la estructura etaria entre los EN y GC. Un 30% del total de casos que forman parte del grupo control tienen edades superiores a los 55 años, mientras que en el grupo de emprendedores este porcentaje es del 9%, no existiendo ningún individuo mayor a 77 años. El ajuste se realiza eliminando del GC todas las personas cuya edad fuese superior a los 77 años y seleccionando al azar una de cada cuatro personas pertenecientes al estrato etario de 55 a 77 años. Finalmente, el estrato de edad superior a 55 años representa el 9,2% dentro del grupo control, consistente con la estructura del grupo de EN.

También se eliminan de la muestra los componentes del mismo hogar en el cual se detecte un emprendedor naciente. El tamaño final de la muestra, ya depurado el GC, se presenta en el siguiente cuadro.

CUADRO 2-4
Tamaño muestral del grupo control por ciudad
luego de la depuración

Ciudad	Tamaño
Córdoba	1.756
Mar del Plata	1.038
Bahía Blanca	957
Rafaela	816
Tandil	493

Fuente: Elaboración propia.

En el punto anterior se presentó la metodología de detección de los EN que conforman el panel, para cada una de las ciudades, y en este punto se explica la determinación del GC. A partir de acá, en el Capítulo 3 se avanza sobre otro de los objetivos del trabajo que es establecer las características de los EN en comparación con el GC tal como se ha determinado anteriormente.

BIBLIOGRAFÍA

- Argañaraz, N.; Olivero, M. y Garzón, J. (1999), *Córdoba productiva: la importancia de las PYMES*. Córdoba: IRREAL-UIC.
- Bolsa de Comercio de Córdoba (1999), *El balance de la economía argentina en 1998*. Córdoba: Eudecor.
- Dunne, T.; Roberts, M. & Samuelson, L. (1988), "Patterns of firms entry and exit in U.S. manufacturing industries". *RAND Journal of Economics*, 19 (4): 495-515.
- Fothergill, S. & Gudgin, G. (1982), *Unequal growth: Urban and regional employment change in the United Kingdom*. London: Heinemann.
- García Reche, A. (2003), *Política económica sectorial y estructural*. Valencia, España: Tirant lo Blanch.
- Glasmeier, A. (1988), "Factors governing the development of high tech industry agglomerations: A tale of three cities". *Regional Studies*, 22 (4): 287-301.
- Graña, F. (2000), *Creación de empresas: Factores asociados al éxito emprendedor en cinco ciudades argentinas*. Mar del Plata: Martín.
- INDEC (1995), Encuesta Permanente de Hogares, Manual de Instrucciones, Cuestionario Individual.
- Mason, C. (1991), "Spatial variations in enterprise: The geography of new firm formation". In Borrows, R. (ed.) *Deciphering the enterprise culture: Entrepreneurship, petty capitalism and the restructuring of Britain*. Ch. 5: 74-176. London: Routledge.
- Motta, J. y Zavaleta, L. (2004), "Factores de competitividad en la industria autopartista de Córdoba". *9ª Reunión Anual Red Pymes Mercosur*, UNGS. Publicado en Actas.
- Municipalidad de Rafaela, CFI, IPEC, Ministerio de la Producción de Santa Fe (1997), *Relevamiento del Sector Productivo 1996*.
- Plan Estratégico de Rafaela. www.rafaela.gov.ar/per.
- Reynolds, P. & Miller, B (1992), "New firm gestation: Conception, birth and implications for research". *Journal of Business Venturing*, 7.

- Reynolds, P.; Miller, B. & Maki, W. (1993), "Regional characteristics affecting business volatility in the United States 1980". In Karlsson, C.; Johannisson, B. & Storey, D. (eds.) *Small business dynamics*, pp. 78-116. London: Routledge.
- Reynolds, P. (1997), "Who starts new firms? – Preliminary explorations of firms in gestation". *Small Business Economics*, 9: 449-462.