
**Rediseño estadístico de la Encuesta Mensual sobre Empresas Comerciales,
cambio de año base, cambio de clasificador económico SCIAN 2007 y
Retropolación de la serie de tiempo.**

Antecedentes

La Encuesta Mensual sobre Establecimientos Comerciales (EMEC) inició operaciones en 1983 teniendo como objetivo generar información estadística de la actividad comercial en el corto plazo, a nivel nacional y por entidad federativa, sin embargo, los problemas técnicos y de recursos limitaron el objetivo a la producción de indicadores referidos a las áreas metropolitanas de las ciudades de México, Monterrey, NL y Guadalajara, Jal., situación que prevaleció por diez años.

En 1993 se amplía la cobertura del proyecto a 15 ciudades y al año siguiente se incorporan 18 más, así se produjeron indicadores mensuales a los cuales se añadió en 1998, tres ciudades más, y en esa misma época se incorpora Durango, Dgo., para llegar a un total de 37 ciudades objeto de estudio cuya serie culmina en 2014.

A pesar de que el estudio de ciudades y áreas urbanas y metropolitanas se justificaba en virtud de que la actividad comercial se desarrolla en función a la población consumidora, era necesario cubrir todo el territorio nacional, así mismo se requería que la encuesta contribuyera a la construcción del Producto Interno Bruto Estatal, por lo que resultaba imperioso incluir en la EMEC muestra de los estados de Hidalgo, Nayarit y Tlaxcala, que no estaban representados por ninguna ciudad.

Así mismo, la EMEC generaba números índice base 2003=100, lo cual desde el punto de vista ortodoxo requería el cambio de año base, pues la referencia a dicho año ya no es cercana para lo que los usuarios de la información pueden percibir.

Adicional a todo lo anterior, el diseño estadístico de la EMEC había sido, por razones técnicas “No probabilístico”, enfocándose a la selección de las empresas grandes, lo cual favorecía la cobertura de la encuesta, no obstante en algunos dominios de estudio las unidades pequeñas no estaban representadas.

Cambios a la Encuesta Mensual sobre Empresas Comerciales (EMEC)

Derivado del panorama que se presentaba en la EMEC, en 2011 se inició la planeación de los cambios del proyecto, a saber:

- Cambio de diseño estadístico
- Cambio de año base
- Adopción del SCIAN 2007
- Ampliación de la cobertura geográfica
- Retropolación de la serie de tiempo

Para consolidar los cambios, se harían pruebas a través de la ejecución en la Encuesta Anual del Comercio del año 2012 y, paulatinamente se trasladarían bajo un riguroso análisis las mejoras encontradas a la Encuesta Mensual sobre Empresas Comerciales.

Cambio de diseño estadístico

La EMEC se realizaba a partir de un diseño estadístico No probabilístico por efectos técnicos y de economía. La ausencia de mecanismos para la actualización del marco de muestreo del sector de actividad más populoso de la economía impedía la consolidación de un diseño estadístico Probabilístico, así en 2012 y gracias a los trabajos del Directorio Estadístico Nacional de Unidades Económicas (DENUE), que logró capitalizar actualizaciones de directorios empresariales provenientes de las mismas encuestas económicas, de convenios con el Sistema de Administración Tributaria (SAT), Instituto Mexicano del Seguro Social (IMSS), Comisión Federal de Electricidad (CFE), etc., ya se contaba con un marco de empresas comerciales aceptablemente actualizado para utilizar un esquema de muestreo que no se deteriorara tanto con el paso de los años.

No es materia de este documento detallar el diseño estadístico, simplemente pondremos en contexto la dimensión del cambio.

El diseño estadístico previo enfocaba la selección de las unidades de muestra por la cobertura del Sector de actividad comercial para cada ciudad objeto de estudio, mientras que el nuevo diseño se abocó a, primero, la selección de Ramas de estudio que podían mantenerse con un esquema No probabilístico debido a la gran concentración de ingresos y personal ocupado en pocas empresas a nivel nacional, y; segundo, las Ramas que por el número de establecimientos a nivel nacional y que por su tamaño en las variables de interés (Personal ocupado e Ingresos por suministro de bienes y servicios) es más adecuada la estimación de su volumen por el método Probabilístico.

Hay que mencionar, que la producción de la estadística de comercio no puede interrumpirse, por lo que todas las acciones se llevaron a cabo en paralelo y, para enfatizar este hecho es relevante anotar que la muestra de empresas comerciales que estaba generando los indicadores, a pesar de una alta coincidencia con la nueva muestra (4,064 de empresas comunes), sufrió cambios importantes al

incorporar 3,259 empresas nuevas y proponer la salida de la muestra de 3,550 unidades.

- El **tamaño de muestra aproximado resultó en 7,323ue**, incorporándose **3,259ue nuevas**; estas últimas serán verificadas en campo durante su levantamiento.
- De la misma forma **causarían baja de la muestra 2011, alrededor de 3,550ue**.
- En enero 2012 se integran **1,083ue**

Cambio de año base

Tradicionalmente las Encuestas Económicas Nacionales del INEGI buscan dar continuidad a los hallazgos de los Censos Económicos, facilitando con esto la elaboración de la contabilidad nacional, de esta manera la elección del año base de la EMEC debía estar en función a los otros proyectos mencionados, siendo 2008 el periodo más adecuado.

Adopción del SCIAN 2007

La EMEC utilizaba en 2012 el Sistema de Clasificación Industrial de América del Norte (SCIAN) 2002, por lo que entre los cambios naturales estaba también el relevo del clasificador a su versión 2007, pues éste último suprimía una Rama y sus correspondientes Clases de actividad comercial, amén de ajustes al interior de algunas Ramas.

Ampliación de la cobertura geográfica

El cambio más radical del proyecto obedecía a la necesidad de contribuir al cálculo del Producto Interno Bruto (PIB) de cada entidad federativa, por lo que el diseño estadístico también enfocó este reto y generó en paralelo indicadores para todas las entidades, por sector de actividad comercial al por mayor y al por menor.

Retropolación de la serie de tiempo

En paralelo al análisis y puesta en marcha de cada uno de los cambios de la encuesta, se mantuvo en la agenda el tema de la cantidad de información disponible así como la no disponible para concretar el nuevo proyecto de la Encuesta Mensual sobre Empresas Comerciales.

Es importante recordar, que las estadísticas económicas sirven a una multitud de propósitos, para lo cual es relevante presentar series de tiempo de al menos cinco años, entre otras razones, porque las series originales requieren un proceso adicional de ajuste estacional (o desestacionalización, que en el caso de encuestas mensuales requiere de 60 periodos para poder efectuar el ajuste mediante el modelo ARIMA), de esta manera, situados en 2012, debíamos definir procedimientos para retropolar información a nivel de microdatos o registros de muestra al menos hasta el año 2008, año seleccionado como base de los índices de actividad comercial.

La Encuesta Mensual sobre Establecimientos Comerciales ya recogía desde 1993 información nacional de las empresas comerciales, así como por ciudad objeto de estudio, pero con el cambio de corte geográfico esos datos ahora debían ser transformados a información por entidad federativa.

En la misma frecuencia de la cobertura geográfica, la problemática se acentuaba en los casos de empresas que ya reportaban ciudades en las que concurrían dos entidades:

- Área metropolitana de la Ciudad de México, a la que había que separar al Distrito Federal la zona conurbada del Estado de México.
- Comarca Lagunera, conformada por Gómez Palacio y Lerdo, Durango y Torreón, Coahuila.
- Tampico Madero, conformada por Río Blanco, Veracruz y los municipios de Tampico y Madero, Tamaulipas.
- Puebla, cuya área metropolitana incluía algunos municipios de Tlaxcala

A los casos antes descritos se suman los de las empresas que participan en dos o más ramas de actividad comercial, pues el cálculo o desglose de su información abarca un registro estadístico de proceso por cada rama y ciudad objeto de estudio. En la EMEC tenemos empresas con actividad hasta en tres ramas distintas.

Así mismo, había unidades de muestra que se incorporaban en 2012 y su primera participación en las encuestas sería a través de la Encuesta Anual del Comercio que recogía información de 2011; éstas mismas unidades habían sido seleccionadas del marco del Censo Económico 2009, por lo que se disponía de información censal del año 2008, pero teníamos el faltante de los años 2009 y 2010, esto es 24 meses que debíamos interpolar.

Una herramienta que resultó muy útil al momento del prorratio o distribución de información empresarial nacional anual en información mensual por entidad federativa, fue la Relación de Unidades Registradas en un Cuestionario, el formato

RURC, en el cual se vierte para cada uno de los establecimientos de la empresa los valores de personal ocupado total y el porcentaje de ingresos que aporta al total empresarial en un año determinado, su Clase de actividad comercial (código de seis dígitos del SCIAN), además del año de inicio de operaciones y su domicilio completo conforme a lo establecido en la Norma Técnica de Domicilios Geográficos (NTDG).

La información anualizada de cada registro (CE 2009 o EAC 2012 o interpolación anualizada), se distribuiría para cada mes conforme al comportamiento observado en los resultados de la Encuesta Mensual sobre Establecimientos Comerciales, según su Rama de actividad comercial en la ciudad con mayor similitud (criterio conocido también como del vecino más próximo o conglomerado).

Así, la interpolación de los valores mensuales de 2009 y 2010, también tenían que reconocer la evolución de las empresas nacionales o multiestablecimientos, es decir, se partía del número de puntos de venta en el periodo censal y no podía rebasarse el número de puntos de venta reportados para la EAC 2012, a menos que el formato RURC tuviera evidencia de nacimientos y bajas de establecimientos que fundamentara un número mayor (o menor) al reportado en 2011.

De esta forma se diseñaron y programaron algoritmos que distribuyeran información con base en razones y proporciones, los que eventualmente se ajustaban en hojas de cálculo.

De entre la gran cantidad de casos de empresas que requirieron estimación de datos faltantes, destacan los que en seguida se enlistan:

- Retropolación de información anualizada del periodo 2012 hacia 2011.
- Interpolación de información anualizada de los periodos 2009 y 2010.
- Distribuciones porcentuales de información anualizada (censos y encuestas), para conseguir información mensual.
- Razones y proporciones para asignar información Empresarial a información por entidad federativa, por variable.
- Cálculo del volumen de variables con base en promedios (de periodos cercanos del mismo registro o de resultados EMEC) de variables relacionadas.
- Distribución de información empresarial por afijación proporcional de entidades federativas.

En seguida se muestran los algoritmos utilizados para los cálculos de microdatos:

EXTRAPOLACIÓN DE MICRODATOS MENSUALES POR DISTRIBUCIÓN PORCENTUAL, SEGÚN RAMA DE ACTIVIDAD COMERCIAL EMEC*

$$Vi_{ma} = \frac{V_{maR}}{V_{aR}} * Vi_a$$

Donde:

V=Valor de cualquier variable

i= Cualquier registro

m= mes

a= año

R= Rama SCIAN

INTERPOLACIÓN DE MICRODATOS ANUALES POR REGISTRO DE PROCESO

$$Vi_a = \frac{(Vi_{a-1}) + (Vi_{a+1})}{2}$$

Donde:

V=Valor de cualquier variable

i= Cualquier registro

m= mes

a= año

RETROPOLACIÓN DE MICRODATOS MENSUALES CON VALORES ABSOLUTOS DE RAMA DE ACTIVIDAD COMERCIAL EMEC*

$$Vi_{m-1a} = \frac{V_{m-1aR}}{V_{maR}} * Vi_{ma}$$

Donde:

V=Valor de cualquier variable

i= Cualquier registro

m= mes

a= año

R= Rama SCIAN

RETROPOLACIÓN DE MICRODATOS MONETARIOS MENSUALES CON FACTOR DE CAMBIO DEL INDICE NACIONAL DE PRECIOS AL CONSUMIDOR

$$Vi_{m-1a} = \frac{INPC_{m-1a}}{INPC_{ma}} * Vi_{ma}$$

Donde:

V=Valor de cualquier variable

i= Cualquier registro

m= mes

a= año

INPC= Índice Nacional de Precios al Consumidor

INTERPOLACIÓN DE MICRODATOS ANUALES CON VALORES ABSOLUTOS DE RAMA DE ACTIVIDAD COMERCIAL EMEC*

$$Vi_a = \frac{V_{a+nR}}{V_{a-1R}} * Vi_{a-1}$$

Donde:

V=Valor de cualquier variable

i= Cualquier registro

n= Periodos anuales

a= año

R= Rama SCIAN

EXTRAPOLACIÓN DE MICRODATOS MONETARIOS MENSUALES CON EL FACTOR DE CAMBIO DEL INDICE NACIONAL DE PRECIOS AL CONSUMIDOR

$$Vi_{ma} = \frac{INPC_{ma}}{INPC_{m-1a}} * Vi_{m-1a}$$

Donde:

V=Valor de cualquier variable

i= Cualquier registro

m= mes

a= año

INPC= Índice Nacional de Precios al Consumidor

CÁLCULO DE REMUNERACIONES TOTALES CON PROMEDIOS DE REMUNERACIONES POR PERSONA DE RAMA DE ACTIVIDAD COMERCIAL EMEC*

$$Rt_{i_{ma}} = Pp_{maR} * Pi_{ma}$$

Donde:

Rt= Remuneraciones totales

i= Cualquier registro

m= mes

a= año

Pp= Promedio de pago mensual por persona ocupada

P= Personal Ocupado mensual

R= Rama SCIAN

EXTRAPOLACIÓN DE MICRODATOS ESTATALES ANUALES POR AFIJACIÓN PROPORCIONAL DE REGISTRO DE PROCESO

$$Vi_{aE} = \left(Vi_{a-1E} / \sum_{E=1}^{32} Vi_{a-1E} \right) * \sum_{E=1}^{32} Vi_{aE}$$

Donde:

V=Valor de cualquier variable

i= Cualquier registro

m= mes

a= año

E= cualquier entidad federativa

Concluidos dichos procesos, se validaba toda la información en el Sistema de Tratamiento Empresarial de la EMEC, el cual consta de 73 validaciones definidas por categoría: sumas (suma de los parciales de entidad igual al total empresarial y, suma de los parciales de conceptos contra totales de temas de análisis), congruencia interna entre variables, congruencia de serie de tiempo o tendencias para registros de proceso nacional y por entidad federativa, etcétera.

TIPO DE VALIDACIÓN SISTEMA EMEC EMPRESARIAL	EMPRESA	FCE	TOTAL
CONGRUENCIA	31	14	45
RELACIONES ANALÍTICAS	3	2	5
SUMAS	9	0	9
TENDENCIA	7	5	12
REMUNERACIONES TOTALES AL PERSONAL OCUPADO			
TOTAL	1	1	2
TOTAL	51	22	73

Adicionalmente hay que mencionar, que durante el lapso de 2012 a 2014 no se interrumpió la generación y divulgación de estadísticas del comercio, gracias a que la transición en la captación de información ocurrió en dos órdenes:

1. Transformación de información de ciudad a entidad, con captación por ciudad.
2. Transformación de información de entidad a ciudad, con captación por entidad.

Para lograr lo anterior, las fuentes informantes recibieron un escrito destacando los cambios de la encuesta y el periodo específico de cuándo les correspondía modificar la entrega de datos de ciudad a datos por entidad federativa; esta estrategia a su vez nos permitió probar y efectuar gradualmente los ajustes a los sistemas de captación por internet, de tratamiento de la información y de proceso de datos.

Adicionalmente, la transición implicó dos procesos paralelos de generación de información estadística, lo que también nos permitió evaluar la congruencia de los resultados obtenidos y ponderar los beneficios generados por los cambios.

El siguiente esquema ilustra de manera resumida los ajustes según tipo de muestra, tipo de empresa y de proceso implicado para lograr la conversión de información captada por ciudad objeto de estudio a información por entidad federativa.

Transición por ciudad a entidad	Periodo	2008		2009 y 2010		2011		2012		2013	
	Proyecto	EMEC	EAC	EMEC	EAC	EMEC	EAC	EMEC	EAC	EMEC	EAC
	Tipo de Empresa / Proceso	Conversión a Entidad		Conversión a Entidad		Conversión a Entidad		Conversión y Captación		Captación	
Muestra coincidente	Empresas uniestablishcimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento
	Empresas multiestablishcimiento y una Rama de actividad	Empresa y Ciudades	Empresa y Establecimientos (RURC)	Empresa y Ciudades	Empresa y Establecimientos (RURC)	Empresa y Ciudades	Empresa y Establecimientos (RURC)	Empresa y Ciudades y Entidades	Empresa y Establecimientos (RURC)	Empresa y Entidades	Empresa y Establecimientos (RURC)
	Empresas multiestablishcimiento y dos o más Ramas de actividad	Empresa y Ciudades	Empresa y Establecimientos (RURC)	Empresa y Ciudades	Empresa y Establecimientos (RURC)	Empresa y Ciudades	Empresa y Establecimientos (RURC)	Empresa y Ciudades y Entidades	Empresa y Establecimientos (RURC)	Empresa y Entidades	Empresa y Establecimientos (RURC)
Fuente de información		Censo Económico 2009		Interpolación		Retropolación EAC		Captación		Captación	
Muestra nueva	Empresas uniestablishcimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento	Establecimiento
	Empresas multiestablishcimiento y una Rama de actividad	Establecimientos, Empresa y Entidades		Empresa y Entidades		Empresa y Entidades	Empresa y Establecimientos (RURC)	Empresa y Entidades	Empresa y Establecimientos (RURC)	Empresa y Entidades	Empresa y Establecimientos (RURC)
	Empresas multiestablishcimiento y dos o más Ramas de actividad	Establecimientos, Empresa y Entidades		Empresa y Entidades		Empresa y Entidades	Empresa y Establecimientos (RURC)	Empresa y Entidades	Empresa y Establecimientos (RURC)	Empresa y Entidades	Empresa y Establecimientos (RURC)

El volumen de información reconvertida fue considerable; abajo se anota el número de registros tratados por mes en el año 2012, que fue cuando se aplicó el último ajuste al diseño estadístico (los números correspondientes hacia 2008 y hacia 2014 son muy similares, aunque diferentes por los efectos naturales de altas y bajas de las empresas en muestra) :

- Muestra de la Encuesta Mensual sobre Establecimientos Comerciales 2012:
 - 8,594 Empresas.
 - 39,273 Establecimientos.
 - 15,414 Firmas Comercial Ciudad (definida como el total de establecimientos de una misma razón social, de una misma Rama de actividad SCIAN 2002, en una misma ciudad objeto de estudio).
- Muestra de la Encuesta Mensual sobre Empresas Comerciales 2012:
 - 7,323 Empresas.
 - 85,894 Establecimientos.
 - 16,566 Firmas Comercial Entidad (definida como el total de establecimientos de una misma razón social, de una misma Rama de actividad SCIAN 2007, en una misma entidad federativa).

Como es posible apreciar, el nuevo diseño estadístico es más eficiente, porque además de aprovechar toda la información de las empresas (no se limita a las ciudades objeto de estudio), con menos unidades de observación capta más del doble de establecimientos del diseño estadístico anterior. Además es relevante hacer mención, que para los dominios de estudio probabilísticos (20 Ramas de

actividad comercial (CIAN 2007), es posible efectuar estimaciones para el total, así como el cálculo de precisiones estadísticas, mismos que se publican mensualmente para mejor comprensión de los usuarios.

Cambios en el instrumento de recolección.

Capítulo VII

Encuesta Mensual sobre Establecimientos Comerciales

VII. FIRMA COMERCIAL CIUDAD Primera parte

CIUDAD OBJETO DE ESTUDIO <small>(El municipio que comprende el territorio que se encuentra dentro de los límites de la ciudad y que incluye al municipio adscrito correspondiente con sus respectivos establecimientos comerciales y que incluye la zona libre de actividad)</small>	NÚMERO DE ESTABLECIMIENTOS		PERSONAL		PERSONERAS QUE TRABAJAN EN LA CIUDAD <small>(Días de trabajo)</small>
	AUTÓNOMOS	CONTRATADOS	DEPENDIENTE DE LA RAZÓN SOCIAL	NO DEPENDIENTE DE LA RAZÓN SOCIAL	
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV

VII. FIRMA COMERCIAL CIUDAD Segunda parte y última

CIUDAD OBJETO DE ESTUDIO <small>(El municipio que comprende el territorio que se encuentra dentro de los límites de la ciudad y que incluye al municipio adscrito correspondiente con sus respectivos establecimientos comerciales y que incluye la zona libre de actividad)</small>	MERCANCIAS COMERCIALES PARA LA REVENTA <small>(Días de trabajo)</small>	INGRESOS		
		POR VENTAS NETAS DE MERCANCIAS COMERCIALES <small>(Días de trabajo)</small>	POR CONSUMACIÓN COMÚN <small>(Días de trabajo)</small>	POR OTRAS ACTIVIDADES ECONÓMICAS <small>(Días de trabajo)</small>
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV
Ciudad _____ CV

Encuesta Mensual sobre Empresas Comerciales

VII. FIRMA COMERCIAL ENTIDAD Primera parte

ENTIDAD COMERCIAL ENTIDAD <small>(El municipio que comprende el territorio que se encuentra dentro de los límites de la entidad y que incluye al municipio adscrito correspondiente con sus respectivos establecimientos comerciales y que incluye la zona libre de actividad)</small>	NÚMERO DE ESTABLECIMIENTOS		PERSONAL		PERSONERAS QUE TRABAJAN EN LA ENTIDAD <small>(Días de trabajo)</small>
	AUTÓNOMOS	CONTRATADOS	DEPENDIENTE DE LA RAZÓN SOCIAL	NO DEPENDIENTE DE LA RAZÓN SOCIAL	
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV

VII. FIRMA COMERCIAL ENTIDAD Segunda parte y última

ENTIDAD COMERCIAL ENTIDAD <small>(El municipio que comprende el territorio que se encuentra dentro de los límites de la entidad y que incluye al municipio adscrito correspondiente con sus respectivos establecimientos comerciales y que incluye la zona libre de actividad)</small>	MERCANCIAS COMERCIALES PARA LA REVENTA <small>(Días de trabajo)</small>	INGRESOS		
		POR VENTAS NETAS DE MERCANCIAS COMERCIALES <small>(Días de trabajo)</small>	POR CONSUMACIÓN COMÚN <small>(Días de trabajo)</small>	POR OTRAS ACTIVIDADES ECONÓMICAS <small>(Días de trabajo)</small>
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV
Entidad _____ CV